Multiscale Methods for Elliptic Partial Differential Equations and Related Applications

Thesis by

Chia-Chieh Chu

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2010

(Defended July 17, 2009)

To my parents

Acknowledgments

I am deeply grateful to my advisor, Prof. Thomas Y. Hou, for his continued encouragement and invaluable guidance on my study, research, and career. I have greatly benefited from his mathematical insight and rigor. He pointed out my weaknesses to me and helped me to overcome them. He also encouraged me to attend some important conferences in my research area and provided financial support to make this possible.

Many people apart from my advisor have contributed to my advancement throughout my graduate studies at Caltech. First and utmost, I would like to heartily thank my collaborator, Prof. Ivan Graham. My interactions with him have increased my knowledge and driven me to achieve my best. He was patience with my poor writing and taught me how to make it better.

Also, I would like to thank Prof. Yalchin Efendiev for many precious discussions and useful advice. His enthusiasm has have made our collaboration fruitful and enjoyable. He is truly an amazing colleague and a remarkable friend.

I would like to thank Prof. Thomas Y. Hou, Prof. Oscar P. Bruno, Prof. Jerrold Marsden and Prof. Tim Colonius, for kindly serving on my defense committee and giving me valuable suggestions.

Thanks to my former and current officemates Peter Stobbe and Alex Gittens, who taught me about American culture and helped me to improve my English a lot.

I would like to thank my parents, who helped take care of my daughter at the beginning of my graduate study, encouraged me when I was depressed, and whose unwavering love was, is, and will be my source of inspiration. Special gratitude goes to my wife, Chen-Yin, who has stood by me for the last five years through the ups and the downs. I would not have been able to make it to this point without her support. Finally, I thank my lovely daughter, Chi-Yeh, for those joyful days she has brought to me.

Abstract

Multiscale problems arise in many scientific and engineering disciplines. A typical example is the modelling of flow in a porous medium containing a number of low and high permeability embedded in a matrix. Due to the high degrees of variability and the multiscale nature of formation properties, not only is a complete analysis of these problems extremely difficult, but also numerical solvers require an excessive amount of CPU time and storage. In this thesis, we study multiscale numerical methods for the elliptic equations arising in interface and two-phase flow problems. The model problems we consider are motivated by the multiscale computations of flow and transport of two-phase flow in strongly heterogeneous porous media. Although the analysis is carried out for simplified model problems, it does provide valuable insight in designing accurate multiscale methods for more realistic applications.

In the first part, we introduce a new multiscale finite element method which is able to accurately capture solutions of elliptic interface problems with high contrast coefficients by using only coarse quasiuniform meshes, and without resolving the interfaces. The method is H^1-conforming, and has an optimal convergence rate of $O(h)$ in the energy norm and $O(h^2)$ in the L_2 norm, where h is the mesh diameter and the hidden constants in these estimates are independent of the "contrast" (i.e. ratio of largest to smallest value) of the PDE's coefficients. The new interior boundary conditions depend not only on the contrast of the coefficients, but also on the angles of intersection of the interface with the element edges. We conduct some numerical experiments to confirm the optimal rate of convergence of the proposed method and its independence from the aspect ratio of the coefficients.

In the second part, we propose a flow-based oversampling method where the actual two-phase flow boundary conditions are used to construct oversampling auxiliary functions. Our numerical results show that the flow-based oversampling approach is several times more accurate than the standard oversampling method. A partial

theoretical explanation is provided for these numerical observations.

In the third part, we discuss "metric-based upscaling" for the pressure equation in two-phase flow problem. We show a compensation phenomenon and design a multiscale method for the pressure equation with highly oscillatory permeability.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Multiscale problems appear in many important areas of modern science and engineering: composite materials, porous media, turbulence transport in high Reynolds number flows, atmosphere/ocean science, and finance. These kind of problems are usually characterized by large disparities in spatial and temporal scales which make them very hard to analyze. One common feature is that the relevant length and time scales are not known a priori but are determined by the solution of the problem itself. A direct numerical simulation of the multiple scale problems is very difficult. To resolve a fine scale in simulation requires an extremely large amount of computer storage and CPU time. Traditional monoscale approaches have proven to be inadequate, even with the fastest supercomputers, because of the range of scales and the prohibitively large number of variables involved.

A typical example is *reservoir simulation*. A reservoir is usually *heterogeneous*; its properties heavily depend on the spacial location. The permeability of the rock may vary from one millidarcy (md) to thousands md. Therefore, a complete computation may involve millions of degrees of freedom, which can easily exceed the limit of today's computers. Fortunately, for most applications it is sufficient to capture the behavior of some macroscopic structure accurately, for example, the *oil production rate* in reservoir simulation. This is why multiscale methods are so widely used, and are of great practical interests: they allow us to solve the problem on a coarse scale, with much fewer degrees of freedom, while retaining a satisfactory accuracy. In Section 1.2, we introduce a well-developed multiscale method, multiscale finite element method (MsFEM). The numerical methods developed in this thesis are motivated by the main

idea of MsFEM.

Among all the multiscale problems, the following elliptic equation in divergence form may be the most intensively studied:

$$\begin{aligned} -\nabla \cdot (\alpha \nabla u) &= f \quad \text{in} \quad \Omega, \\ u &= g \quad \text{on} \quad \partial\Omega. \end{aligned} \tag{1.1.1}$$

Because of highly oscillatory coefficient α, a simulataneous resolution of all scales in this problem is impossible computationally. One therefore seeks to exploit scale separation to split the problem into a hierarchy of tractable models. More precisely, one assumes $\alpha(x) = \alpha(x/\varepsilon)$, with ε a small parameter, and α a smooth positive-valued periodic function on a unit cell Y. This can be viewed as an ideal model for infinite perfectly periodic crystal lattice. Under this assumption, the *homogenization theory* [1] has been developed and the concepts of Γ-*, G-, H*-convergence[2] have been introduced. Numerous multiscale schemes have been developed based on this framework. Most of these methods can also be applied to non scale separation cases, but lack theoretical foundation guaranteeing convergence in this cases. Recently, Owhadi and Zhang [78] revealed a new type of compensation phenomena for PDEs with coefficients that do not necessarily satisfy scale separation assumption. They designed numerical homogenization methods by using a coordinate transformation and provided a rigorous mathematical analysis.

In this thesis, we will focus on two types of $\alpha(x)$ in equation (1.1.1): α is piecewise constant with *high contrast*, and $\alpha(x) = \lambda(x)k(x)$, which arise in interface problems and two-phase flow problems respectively. We emphasize that both types of problem do not exhibit scale separation or periodic structures. Hence we cannot follow the traditional approach of applying homogenization theory by exploring scale separation. Using the local properties of the exact solutions of interface problems, we create a special type of MsFEM and prove a robust convergence rate. More detail is given in Section 1.3, 1.4 and later chapters.

This thesis consists of 5 chapters. In Chapter 1, we introduce the governing equa-

[1] A good approximation of the macroscopic property of such materials can be obtained by letting the parameter ε tend to zero. Homogenization theory describes these limit processes. We refer the reader to the books of Bensoussan, Lions, and Papanicolaou [18] and Jikov, Kozlov, and Oleinik [58] for more information.

[2] The Γ-convergence is an abstract notion of functional convergence which was introduced by De Giorgi [44]. *G*-convergence is a notion of convergence associated with sequences of symmetric, second-order, elliptic operators which was introduced by Spagnolo [85]. Murat and Tartar [77] generalized *G*-convergence to *H*-convergence to handle the case of non-symmetric problems.

tions and review a selected portion of the related literature. Chapter 2 discusses high contrast interface elliptic problems; We formulate our numerical methods and demonstrate their robustness. In Chapter 3, we study two-phase flow problems and compare two different strategies of oversampling techniques [3]. We generalize the metric based up-scaling technique of Owhadi and Zhang [78] to two-phase flow problems in Chapter 4. In Chapter 5, we give concluding remarks and discuss possible directions for future work.

1.2 Multiscale Finite Element Method

The multiscale finite element method, in which basis functions are computed by solving local homogeneous PDEs subject to special boundary conditions, has a large literature. The primitive form of this method can be traced back to the early work of Babuška, Caloz and Osborn [14, 12] who introduced special basis functions for 1D elliptic problems with rough coefficients.

In [51], Hou and Wu developed the multiscale finite element method for multidimensional problems with multiscale coefficients. The main idea of the method was to incorporate the small scale features in the underlying physical problem into the finite element bases locally. Within each coarse grid block, one can construct the multiscale basis by solving the leading order governing equation locally. The small scales then interact with the large scales through the variational formulation of the finite element method.

To illustrate the method more clearly, consider the linear elliptic equation (1.1.1). Let $\mathcal{T}_h$ be a partition of Ω into finite elements (triangles, quadrilaterals and so on). For each interior node x_p of the mesh $\mathcal{T}_h$, we define the nodal basis function Φ_p^{MS}, whose restriction $\Phi_{p,\tau}^{\mathrm{MS}}$ to each $\tau \in \mathcal{T}_h$ is required to solve the "*local problem*":

$$-\nabla \cdot (\alpha \nabla \Phi_{p,\tau}^{\mathrm{MS}}) = 0 \quad \text{in} \quad \tau, \tag{1.2.1}$$

together with a suitable boundary condition:

$$\Phi_{p,\tau}^{\mathrm{MS}} = \phi_{p,\tau}, \quad \text{on} \quad \partial\tau \ , \quad \text{with} \quad \phi_{p,\tau}(x_q) = \delta_{p,q}, \quad \text{for all vertices } x_p, x_q \text{ of } \tau \ .$$

[3] We remark that the notion of oversampling we consider here refers to the oversampling in the physical space when we construct a local multiscale basis function, which is different from the oversampling trechique in the frequency space used in signal analysis.

The boundary condition $\phi_{p,\tau}$ is usually chosen to be the restriction of the standard finite element basis functions on $\partial\tau$. In general, one solves (1.2.1) on the fine grid to obtain basis functions. Nevertheless, the basis function can be computed analytically in some cases. We extend the basis functions $\Phi_p^{\rm MS}$ trivially on Ω and define the multiscale finite element space $\mathcal{V}_h^{\rm MS} = {\rm span}\{\Phi_p^{\rm MS}\}$. Then the multiscale finite element method is to find $u_h^{\rm MS} \in \mathcal{V}_h^{\rm MS}$ such that

$$\sum_{\tau\in\mathcal{T}_h}\int_\tau \alpha(x)\nabla u_h^{\rm MS}\cdot\nabla v_h^{\rm MS}dx \;\; = \;\; \int_\Omega f\, v_h^{\rm MS}dx\ , \quad \text{for all } v_h^{\rm MS}\in\mathcal{V}_h^{\rm MS}\ .$$

In 1D, $\partial\tau$ consists of only two points and there is no need to choose boundary condition. This unique phenomena in 1D leads to an interesting superconvergence result, i.e., $u_h^{\rm MS} = u$ at the nodal points, which has been shown by Hou, Wu and Cai [52]. However, this superconvergence result may not be true in two or higher dimensions. Hou and Wu [51] identified a key issue in multiscale methods in multi-dimensions, that is the microscopic boundary condition which connects the small scale bases to the macroscopic solution. To improve the accuracy significantly, they suggested to use the "oscillatory" boundary condition which is defined as the solution of the reduced elliptic equation on $\partial\tau$:

$$\frac{\partial}{\partial s}\alpha_{\,b}(s)\frac{\partial\phi_{p,\tau}}{\partial s} = 0, \tag{1.2.2}$$

where s is arc-length parameter of $\partial\tau$ and $\alpha_{\,b}$ is the restriction of α on $\partial\tau$. They further performed a convergence analysis to reveal a resonance error introduced by the microscopic boundary condition [51, 52]. An oversampling technique was proposed to effectively reduce the resonance error [51]. The multiscale bases are strongly localized and adaptive. In many cases, the multiscale bases can be precomputed and used repeatedly in subsequent computations with different source terms and different boundary conditions. In some situations the basis functions can be updated adaptively. This leads to significant computational saving in upscaling the two-phase flows where the elliptic (pressure) equation needs to be solved many times dynamically. Another attractive feature of the method is the ability to both coarsen (upscale) a fine grid solution and to reconstruct (downscale) the fine grid solution from a coarse grid solution by using the multiscale bases. This property is very attractive in many

engineering applications.

Most of the convergence analyses of multiscale finite element methods are for the periodic homogenization case where, in (1.1.1), $\alpha(x) = \alpha(x/\varepsilon)$, with ε a small parameter, and are geared toward obtaining optimal convergence, robust with respect to the "oscillation parameter" $\varepsilon \to 0$ (e.g., [51, 52]). However, the method itself is quite general and has been applied to non-periodic cases with considerable success. For example, in [51] one finds an application of (1.1.1) to the case when α is a realization of a random field, both in the isotropic and anisotropic cases and with highly contrasting media. We refer the reader to a recent book by Efendiev and Hou [39] for more discussions on the theory of multiscale problems and applications of and related numerical methods.

1.3 Elliptic Interface Problems

Second order elliptic and parabolic equations with discontinuous coefficients– interface problems– are often encountered in material science and fluid dynamics. For instance, they occur when considering a system containing two or more distinct materials with different attributes (conductivity, density, diffusions, etc). The solution of interface problems is an integral component in an even larger collection of applications, including the projection method for solving Navier-Stokes equations with two-phase flow [26, 18] and the Hele-Shaw flow [50, 48].

The coefficients in interface problems typically have a jump across the interface between materials. Due to the jump, the solution loses global regularity,[4] and therefore, standard finite element and finite difference methods fail to achieve high orders of accuracy. Numerous numerical methods have been proposed to explore the jump conditions in order to obtain higher order convergence. We introduce some related work in Section 1.3.2.

1.3.1 Equations and Weak Formulation

In this subsection, we briefly describe the strong and weak formulations for the elliptic interface problem. For simplicity, we assume there is only one interface. Formulas for multiple interfaces can be derived analogously.

[4] See Section 2.2.2 for more discussion.

Let Ω be a bounded domain and Γ be an interface which divides Ω into two disjoint subdomains, Ω^+ and Ω^-. Elliptic interface problems are usually formulated as

$$\begin{aligned} -\nabla \cdot (\alpha \nabla u) &= f \quad \text{in} \quad \Omega, \\ u &= g \quad \text{on} \quad \partial\Omega, \end{aligned} \tag{1.3.1}$$

where the coefficient α is a symmetric, uniformly elliptic matrix whose components are C^1 on each subdomain Ω_i, but may be discontinuous across the interface Γ. Since α is not C^1 on Ω, equation (1.3.1) is only meaningful in the weak sense:

$$\int_\Omega \alpha(x)\nabla u(x) \cdot \nabla v(x) dx \;=\; \int_\Omega f(x)v(x)dx \quad \text{for all} \quad v \in H_0^1(\Omega), \tag{1.3.2}$$

where the trace of u is equal to g on $\partial\Omega$. Using integration by parts, we obtain the jump conditions

$$[u]_\Gamma \;=\; u^+ - u^- = 0, \tag{1.3.3}$$

$$\left[\alpha \frac{\partial u}{\partial \mathbf{n}}\right]_\Gamma \;=\; \alpha^+ \frac{\partial u^+}{\partial \mathbf{n}} - \alpha^- \frac{\partial u^-}{\partial \mathbf{n}} = 0, \tag{1.3.4}$$

where $\mathbf{n}$ is the normal unit vector of Γ, and the "$\pm$" superscripts refer to limiting values on Γ taken from the subdomains $\Omega^\pm$.

Sometimes an interface problem can also be caused by a singularity in the source term f. More precisely, the flux $\alpha\frac{\partial u}{\partial \mathbf{n}}$ is not continuous across the interface Γ when f is of the form

$$f(x) = f_s(x) - \int_\Gamma q(X(s))\delta(x - X(s))ds,$$

where f_s is a smooth function,[5] δ is the Dirac-delta function, $X(s)$ is the arc-length parametrization of the interface, and q is the source strength on the interface. In this case, the jump condition of flux (1.3.4) is modified to

$$\left[\alpha \frac{\partial u}{\partial \mathbf{n}}\right]_\Gamma = \alpha^+ \frac{\partial u^+}{\partial \mathbf{n}} - \alpha^- \frac{\partial u^-}{\partial \mathbf{n}} = q. \tag{1.3.5}$$

[5]Smoothness here means f_s is in $C(\Omega)$ or $L_2(\Omega)$.

Therefore, equation (1.3.1) is often written in the equivalent strong form

$$\begin{aligned} -\nabla \cdot (\alpha \nabla u) = f_s \quad \text{in} \quad \Omega^{\pm}, \quad u = g \quad \text{on } \partial\Omega, \\ [u]_\Gamma = u^+ - u^- = 0, \quad \left[\alpha \frac{\partial u}{\partial \mathbf{n}}\right]_\Gamma = \alpha^+ \frac{\partial u^+}{\partial \mathbf{n}} - \alpha^- \frac{\partial u^-}{\partial \mathbf{n}} = q. \end{aligned} \tag{1.3.6}$$

The reader should keep in mind that our method is designed to match the jump conditions (1.3.3) and (1.3.4).

1.3.2 Literature Review

Before presenting our results, we first discuss relevant existing literature on interface problems. The approach closest to ours is the immersed finite element (IFE) method of Li, Lin and Wu [68]. They considered uniform triangular grids and approximated the interface by a straight line segment when it intersects a coarse grid element. By matching the jump condition, they created a special basis function for elements which are cut through by the interface and proved a second order convergence rate in the L_2 norm and a first order convergence rate in the H^1 semi-norm. On the other hand, their analysis relies on regularity assumptions of the exact solution u and the constants in their error estimates depend strongly on the contrast of the coefficient.[6] By way of contrast, our error estimates do not depend on the contrast of α, and the Sobolev norm of u can be bound by the Sobolev norm of source tern f a priori. Moreover, it turns out that when the interface intersects an element in a straight line, our new method coincides with the method of [68]. The connections between our method and that of [68] are further discussed in Section 2.4.2.

Much earlier, Babuška [11] studied the convergence of methods based on a minimization problem equivalent to (2.1.1) in which the boundary and jump condition were incorporated in the cost functions. The convergence rate of the method was proved under technical assumptions. There are many subsequent works on such penalty methods, for example Barrett and Elliott [16]. Another relevant work is due to Chen and Zou [22] who approximated the smooth interface by a polygon and used classical finite element methods to solve both elliptic and parabolic interface equations. The disadvantage of this approach is that the mesh must align

[6] In [68], Li, Lin and Wu assumed the exact solution u is C^2 on $\Omega^{\pm}$ respectively and interface Γ is C^1. They obtained the error estimate $\|u - u_h\|_{1,\Omega} \leq C_5\, h$. The hidden constant C_5 is bounded by $C\,\widehat{\alpha}^4 \|D^2 u\|_\infty$, where $\widehat{\alpha}$ is the contrast $(\max\alpha/\min\alpha)$ and C is a constant independent of u, α, h.

with the interface. Plum and Wieners [82] studied interface problems with piecewise constant coefficients and proved (under certain assumptions) optimal *a priori* estimates which are independent of the coefficients for standard finite element methods with meshes resolving the interface. Related results for discontinuous Galerkin methods were given by Dryja in [31]. These body fitted methods[7] require to generate complicated meshes to resolve interfaces, and therefore, many efficient and popular packages/solvers[8] which are written in unform grids, cannot be applied directly. Unfitted high order finite element methods were recently studied by Li et. al. in [65], and error estimates which are explicit in both the order of the elements and the error in the boundary approximation were proved.

There has been a lot of effort in developing accurate and efficient finite difference methods for the interface problem. Among them, the Immersed Boundary method (IBM) was developed by Peskin [80] for studying the motion of one or more massless, elastic surfaces immersed in an incompressible, viscous fluid, particularly in biofluid dynamics problems where complex geometries and immersed elastic membranes are present. The IBM method employs a uniform Eulerian grid over the entire domain to describe the velocity field of the fluid and a Lagrangian description for the immersed elastic structure. The interaction between the fluid and the structure is expressed in terms of the spreading and interpolation operations by use of smoothed delta functions. We refer to [81] for an extensive review of this method and its various applications. Motivated by Peskin's method, Unverdi and Tryggvason [89] have developed a highly successful front tracking method to study viscous incompressible multiphase flows.

Another related work is the Immersed Interface Method (IIM) for elliptic interface problems developed by LeVeque and Li [64]. The main idea of the IIM method is to use the jump condition across the interface to modify the finite difference approximation near the interface. When this is done properly, one can achieve a second order discretization. The IIM method can also be applied to the moving interface problem [48] and to the irregular domain problem [32]. Several extensions and improvements can be found in the references [3, 66, 67].

An important development in interface capturing methods is the Ghost Fluid

[7] Body fitted methods mean the meshes of the methods are aligned the interface.
[8] For examples, fast Poisson solvers, Clawpack [63] for conservation laws and FFT packages etc.

Method (GFM) [42] developed by Fedkiw, Merriman, Aslam and Osher. The GFM method incorporates the interface jump condition into the finite difference discretization in a clever way which can be implemented efficiently. The GFM method has been applied with considerable success to capturing discontinuities in multimedium compressible flow [61] and strong shock impacting problems [62]. The GFM method has been generalized to the elliptic interface problem in [71] and its convergence property has been analyzed in [72]. Related works include [93, 21]. The matched interface and boundary method [93, 92] proposed by Zhou et al. can be viewed as a generalization of IIM and GFM. The authors introduced friction points to match the jump condition to very high order. Very recently, Chern and Shu [21] proposed the coupling interface method for the elliptic interface problem. They provided both first order and second order versions of the method. We remark that there has been little progress on the development of rigorous convergence theory for finite difference methods for interface problems. By way of contrast, the present work proves the robust convergence of our proposed method.

1.3.3 Summary of the Results

In Chapter 2 we propose a new type of multiscale finite element methods for the elliptic interface problem:

$$\int_\Omega \alpha(x)\nabla u(x)\cdot\nabla v(x)dx \;=\; \int_\Omega f(x)v(x)dx\;, \quad v\in H_0^1(\Omega)\;, \tag{1.3.7}$$

where α is piecewise constant with respect to the partition $\{\Omega_i : i = 0,\ldots,m\}$ of Ω. Our method involves special "multiscale" nodal basis functions on a (coarse) quasiuniform triangular mesh $\mathcal{T}_h$. On elements on which α is constant, these basis functions just coincide with the usual linear hat functions. Otherwise we pre-compute the basis function by solving a homogeneous version of (1.3.7) on the relevant elements, subject to special boundary conditions described later in Chapter 2. The resulting basis functions are then used to define a multiscale finite element solution u_H^{MS} by the usual Galerkin method.

We show that our method satisfies an error estimate of the form

$$|u-u_h^{\mathrm{MS}}|_{H^1(\Omega)} \;\le\; |u-u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \;\le\; Ch\left[h|f|^2_{H^{1/2}(\Omega)}+\|f\|^2_{L^2(\Omega)}\right]^{1/2}\;, \tag{1.3.8}$$

where

$$|v|^2_{H^1(\Omega),\alpha} \;=\; a(v,v)\;, \quad \text{with} \quad a(v,w) = \int_\Omega \alpha \nabla v \cdot \nabla w\;,$$

and the constant C is independent of h and of the contrast parameter $\widehat{\alpha}$. This should be compared to the best result of $\mathcal{O}(h^{1/2-\varepsilon})$, with a hidden constant which generally depends on the contrast, for standard finite element methods on a mesh which does not resolve the interface. We also devise a non-standard duality argument which shows that

$$\|u - u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \;\leq\; Ch^2 \left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L^2(\Omega)} \right]^{1/2}\;, \tag{1.3.9}$$

The price to pay for this improved convergence rate is the solution of subgrid problems on elements which straddle the interface and a slightly worse dependence than normal on the data f on the right hand sides of (1.3.8) and (1.3.9). The local subgrid problems can be done as a preprocessing step before solution of the global finite element problem. The accuracy needed for these subgrid problems is investigated numerically in Section 2.5.

The convergence analysis of the new multiscale finite element method devised in this work makes no appeal to homogenization theory but nevertheless explains why "multiscale"-type basis construction can be beneficial in more general situations. It turns out that the new interior boundary conditions obtained in the present work are a genuine generalization of the "oscillatory" boundary conditions of [51], in the sense that the two coincide if and only if the interfaces intersect the element edges orthogonally. Some of the arguments used in this paper have already been developed in the context of domain decomposition methods in [45, 46, 79].

To our knowledge the dependence of the accuracy of numerical methods for elliptic interface problems on coefficient contrast has not been previously analyzed, even though such high contrast problems are ubiquitous in porous media problems, especially in geophysical and oil recovery applications. We note, however, that there is substantial literature on the performance of iterative methods (see, e.g. [45] and the references therein) and on the analysis of a posterior error estimates (see, e.g. [19, 4, 90]) for such high contrast problems, but this literature does not address the issues considered in the present work.

1.4 Two-Phase Flow in Porous Media

We now introduce the two-phase flow problem. A porous medium is a solid medium that is permeated by miniscule pores that allow the transport of a gas or a fluid. A petroleum reservoir is a porous medium that contains hydrocarbons. The term "phase" stands for matter that has a homogeneous chemical composition and physical state. Solid, liquid, and gaseous can be distinguished. The primary goal for studying these subjects is to predict the performance of a reservoir and find the ways to optimize recovery of hydrocarbons.

In the very early stage, a reservoir essentially contains a single fluid such as gas or oil, and the pressure is sufficient high enough that the gas or oil is produced by natural decompression without any pumping effort at the wells. By the time a pressure equilibrium between the oil field and the atmosphere occurs, it usually leaves 70%-85% of hydrocarbons in the reservoir. To recover part of the remaining oil, water is injected into injection wells while the oil is produced through the production wells. This process maintains high reservoir pressure and flow rate. Therefore, in reservoir simulation, people are interested in the simultaneous flow of two or more fluid phases within a porous medium.

Flow in a porous medium is governed by the *conservation of mass, Darcy's law* and an *equation of state* which lead to a complicated system of PDEs. For mathematical understanding, in this section, we introduce a simplified but representative model, the *pressure and saturation equations*. We also review some existing upscaling techniques.

1.4.1 Derivation of the Model Equations

We consider two-phase flow in a reservoir Ω under the assumption that the displacement is dominated by viscous effects; i.e., we neglect the effects of gravity, compressibility, and capillary pressure. Porosity is assumed to be constant. In this flow problem, the two phases are taken to be water and oil, designated by subscripts w and o, respectively. We write Darcy's law, with all quantities dimensionless, for each phase as follows:

$$\mathbf{v}_j = -\frac{k_{rj}(S)}{\mu_j}\mathbf{k}\cdot\nabla p, \tag{1.4.1}$$

where $\mathbf{v}_j$ is the phase velocity, $\mathbf{k}$ is the permeability tensor, k_{rj} is the relative permeability to phase j ($j = o, w$), S is the water saturation (volume fraction), p is pressure and μ_j is the viscosity of phase j ($j = o, w$). In this work, a single set of relative permeability curves is used and $\mathbf{k}$ is assumed to be a diagonal tensor. Combining Darcy's law with a statement of conservation of mass allows us to express the governing equations in terms of the so-called pressure and saturation equations:

$$-\nabla \cdot (\lambda(S)\mathbf{k} \cdot \nabla p) = h, \tag{1.4.2}$$

$$\frac{\partial S}{\partial t} + \mathbf{v} \cdot \nabla f(S) = h_w, \tag{1.4.3}$$

where λ is the total mobility, f is the fractional flow of water, $h = h_w + h_o$ is a source/sink term and $\mathbf{v}$ is the total velocity, which are respectively given by:

$$\lambda(S) = \frac{k_{rw}(S)}{\mu_w} + \frac{k_{ro}(S)}{\mu_o}, \quad f(S) = \frac{k_{rw}(S)/\mu_w}{k_{rw}(S)/\mu_w + k_{ro}(S)/\mu_o}, \tag{1.4.4}$$

$$\mathbf{v} = \mathbf{v}_w + \mathbf{v}_o = -\lambda(S)\mathbf{k} \cdot \nabla p. \tag{1.4.5}$$

The above description is referred to as the fine model of the two-phase flow problem and we look for solutions of S, p as functions of x, t. Typical boundary conditions for (1.4.2) considered in this work are fixed pressure at some portions of the boundary and no-flow conditions on the remainder of the boundary. For the saturation equation (1.4.3), we impose $S = 1$ on the inflow boundaries. For simplicity, in further analysis we will assume that $h_w = h_o = 0$ so that $h = 0$.

1.4.2 Upscaling Procedures

The upscaling of two-phase flow systems has been discussed by many authors [27, 15, 34]. In most upscaling procedures, the coarse-scale pressure equation is of the same form as the fine-scale equation (1.4.2), but with an equivalent grid block permeability tensor $\mathbf{k}^*$ replacing $\mathbf{k}$. For a given coarse-scale grid block, the tensor $\mathbf{k}^*$ is generally computed through the solution of the pressure equation over the local fine-scale region corresponding to the particular coarse block [33]. The coarse-grid $\mathbf{k}^*$ computed in this manner has been shown to provide accurate solutions to the coarse-grid pressure equation. For channelized porous media, the global information can be used in the calculation of effective coarse-grid permeability [24], but these upscaling approaches

are not exact at the initial time.

The high degree of variability and multiscale nature of formation properties such as permeability pose significant challenges for subsurface flow modeling. Geological characterizations that capture these effects are typically developed at scales that are too fine for direct flow simulation, so further work is required to enable the solution of flow problems in practice. Upscaling procedures have been commonly applied for this purpose and are effective in many cases.[9] More recently, a number of multiscale finite element (e.g., [51, 25, 9, 1, 2, 37]) and finite volume [55, 56] approaches have been developed and successfully applied for problems of this type.

Our purpose in this work is to propose a new oversampling strategy for constructing multiscale basis functions within the framework of multiscale finite element method (MsFEM). The oversampling strategy was first proposed by Hou and Wu in [51]. We describe the details of the method in Section 3.2. Besides it, there are a number of multiscale numerical methods (or frameworks) with a similar general objective, such as generalized finite element methods [12], residual free bubbles [84], the variational multiscale method [54], the multiscale finite element method (MsFEM) [51], two-scale finite element methods [73], two-scale conservative subgrid approaches [9], the heterogeneous multiscale method (HMM) [36], and multiscale mortar methods [10]. We remark that special basis functions in finite element methods have been used earlier in [14]. Multiscale finite element methodology has been modified and successfully applied to two-phase flow simulations in [55, 56] and later in [25, 1].

Most multiscale methods presented to date have applied local calculations for the determination of basis functions (or, in the case of variational multiscale methods [9], subgrid integrals). Though effective in many cases, the accuracy of these local calculations may deteriorate for problems in which global effects are crucial. The importance of global information has been illustrated within the context of upscaling procedures in recent investigations [24, 23]. These studies have shown that the use of global information in the calculation of the upscaled parameters can significantly improve the accuracy of the resulting coarse model.

[9] See [91, 83, 43] for reviews and discussion.

1.4.3 Summary of the Results

In Chapter 3, we propose a flow-based oversampling method to compute pressure equation in the two-phase flow problem. The main idea of oversampling techniques is to use solutions of the underlying single-phase flow equation in larger domains to compute the basis functions. These basis functions are used in the two-phase flow simulations with varying (dynamic) mobility. Oversampling techniques reduce the effect of the artificial boundary conditions that are often imposed when computing local quantities, such as upscaled permeabilities or basis functions. When there is no scale separation, the oversampling region is taken to be the entire domain. Typically, generic boundary conditions are used to compute the auxiliary oversampling functions. These boundary conditions do not reflect the actual two-phase flow boundary conditions, which can have large effects, in the simulations. In particular, when two-phase flow boundary conditions contain some types of singularities, the single-phase flow solutions obtained using generic boundary conditions are not sufficient to represent these effects. For this reason, one needs to incorporate the actual two-phase flow boundary conditions. In the proposed flow-based oversampling method, we take one (or more) auxiliary oversampling functions to be the solution of single-phase flow equations with the original (two-phase flow) boundary information. We present a partial analysis which demonstrates the importance of using the actual boundary conditions. Moreover, our analysis explains when one needs to use the actual two-phase flow boundary conditions which are associated with the "singularity" in the boundary conditions of two-phase flows.

To illustrate the performance of this new strategy, we present several representative numerical results. Precisely, comparison between the flow-based and standard oversampling are given for typical two-phase flow and transport simulations. In our numerical experiments, we use the permeability fields from the SPE comparative project [28]. These permeability fields are channelized and difficult to upscale. Due to the channelized nature of these permeability fields, non-local effects are important so often some type of limited global information is used in multiscale simulations (e.g., [37, 35]). In our simulations, we test various viscosity ratios and compare integrated quantities, such as oil production rate and total flow rate, as well as saturation errors at some time instances. In all cases, we observe that the flow-based oversampling

methods are more accurate; in fact, in almost all the cases we consider, it gives several orders of improvement.

In Chapter 4 we consider a multiscale method for the pressure equation in two-phase flow problems:

$$\begin{aligned} -\nabla \cdot (\lambda(\mathbf{x})\mathbf{k}(\mathbf{x}) \cdot \nabla p(\mathbf{x})) &= h \quad \text{in } \Omega, \\ p &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{1.4.6}$$

We assume the mobility λ is a positive C^1-function and h is a L_2 function on Ω. Our method is based on *metric based upscaling* introduced by Owhadi and Zhang [78].

In [78], Owhadi and Zhang developed numerical homogenization methods for divergence form elliptic equations

$$\begin{aligned} -\nabla \cdot (\mathbf{k}\nabla p) &= f \quad \text{in} \quad \Omega, \\ p &= 0 \quad \text{on} \quad \partial\Omega, \end{aligned} \tag{1.4.7}$$

where $\mathbf{k}$ is a $n \times n$ symmetric matrix with entries in $L_\infty(\Omega)$ and f is a $L_r(\Omega)$ function with $r \geq 2$. They discovered that the solution p of (1.4.7) is $W^{2,r}$ with respect to the harmonic coordinates, although it is only $W^{1,r}$ in Euclidean coordinates. The harmonic coordinates $F(\mathbf{x}) = (F_1(\mathbf{x}), ..., F_n(\mathbf{x}))$ associated to (1.4.7) are given by

$$\begin{aligned} \nabla \cdot (\mathbf{k}\nabla F_i) &= 0, \\ F_i(\mathbf{x}) &= x_i, \end{aligned} \tag{1.4.8}$$

which are always automorphisms in 2D [7]. The mechanism of gaining extra differentiability of p is following: $p \circ F^{-1}$ satisfies a non-divergence elliptic equation which is known to have $W^{2,r}$ estimate under a Cordes type condition [74]. More precisely, if $\sigma = (\nabla F)^t \mathbf{k} \nabla F$ is stable,[10] then

$$\|p \circ F^{-1}\|_{W^{2,r}(\Omega)} \leq C \|f\|_{L_r(\Omega)}, \tag{1.4.9}$$

for some constant C independent of u. This compensation phenomena has also been observed numerically, and is important in designing effective numerical homogeniza-

[10] σ is stable if and only if $\beta_\sigma := \text{esssup}_{x \in \Omega} \, (n - (\text{Trace}[\sigma])^2 / \text{Trace}[\sigma^t \sigma]) < 1$ and $\|(\text{Trace}[\sigma])^{n/2r-1}\|_{L_\infty(\Omega)} < \infty$. See [78] for more detail.

tion methods. For example, Owhadi and Zhang used the composition rule to construct the numerical homogenization method. They defined the multiscale finite element space $\mathcal{V}_h$ by

$$\mathcal{V}_h = \text{span}\{\varphi \circ F : \varphi \in X_h\}, \tag{1.4.10}$$

where X_h is a standard C^0 or C^1 finite element space, and proved the error estimate:

$$\|p - p_h\|_{H^1(\Omega)} \leq C\, h \|f\|_{L_r(\Omega)}, \tag{1.4.11}$$

However, this method is only attractive if one needs to solve (1.4.7) multiple times with different source tern f, since it requires solving (4.1.2) n times to obtain the harmonic coordinates F, which is the same cost of solving (1.4.7) n times.[11]

In this work, we show that the compensation phenomena is still true for the pressure equation (1.4.6) with a C^1-function λ by using the harmonic coordinate F independent of λ. We prove that if p is the solution of (1.4.6) and F is given in (4.1.2), both (1.4.9) and (1.4.11) hold under the same assumptions. In practice, we can precompute the metric based upscaling basis functions, and use them to solve (1.4.2) for different source tern f and mobility function λ repeatedly.

[11] It is necessary for *metric based upscaling* method to have harmonic coordinate very accurately. Therefore, one must solve (4.1.2) in fine grid.

Chapter 2

High-Contrast Elliptic Interface Problems

The work presented in this chapter consists of materials from the recent paper [29] by Chu, Graham and Hou.

2.1 Introduction

In this chapter, we present a new application of multiscale finite element methods to the classical elliptic problem in weak form

$$\int_\Omega \mathcal{A}(x)\nabla u(x)\cdot\nabla v(x)dx \;=\; \int_\Omega F(x)v(x)dx\ , \quad v\in H_0^1(\Omega)\ , \tag{2.1.1}$$

where the solution $u\in H^1(\Omega)$ is required to satisfy a Dirichlet condition on $\partial\Omega$ and F is given, on a bounded domain $\Omega\subset\mathbb{R}^2$. To concentrate on the essential aspects of this new theory we treat primarily the homogeneous Dirichlet problem when the boundary of Ω is a convex polygon.[1]

The coefficient $\mathcal{A}$, which is assumed here to be scalar, will be allowed to jump across a number of smooth interior interfaces and the aim of the present work is to propose and analyze a new multiscale finite element method for this problem on coarse meshes which are not required to resolve the interfaces. Our method has the same rate of convergence (with respect to mesh diameter) as is known for the case when $\mathcal{A}$ is globally smooth[2] and, moreover, this rate of convergence is independent of the range of variation ("contrast") of the coefficient function $\mathcal{A}$.

[1] These are not essential restrictions: All our results are true for smooth boundaries as well. Similar results could be obtained for non-convex polygons, treated with local mesh refinement. Moreover we state the corresponding results for the boundary condition $g\neq 0$ later in the chapter.

[2] The rate is not degraded by the loss of solution regularity across the interfaces.

While the method we propose could be used for the case when the interfaces in $\mathcal{A}$ are of general geometry, our proofs are given for the particular case when Ω contains a finite number of inclusions, each with smooth closed boundary not intersecting $\partial\Omega$, such as is depicted in Figure 2.1. Denoting the inclusions by $\Omega_1, \ldots, \Omega_m$ and setting $\Omega_0 = \Omega \backslash \cup_{i=1}^m \Omega_i$, we assume that the coefficient $\mathcal{A}$ is piecewise constant with respect to the decomposition $\{\Omega_i : i = 0, \ldots, m\}$. Again, there is no essential difficulty in generalizing to piecewise smooth coefficients. Setting $\mathcal{A}_{\min} = \min\{\mathcal{A}|_{\Omega_i} : i = 1, \ldots, m\}$, we first scale problem (2.1.1) by dividing by $\mathcal{A}_{\min}$, yielding the weak form: find u such that $u \in H_0^1(\Omega)$ and

$$a(u, v) = (f, v)_{L_2(\Omega)} \quad \text{for all} \quad v \in H_0^1(\Omega) \ , \tag{2.1.2}$$

where

$$a(u, v) = \int_\Omega \alpha(x) \nabla u(x) \cdot \nabla v(x) dx \ , \tag{2.1.3}$$

with

$$\alpha(x) = \frac{1}{\mathcal{A}_{\min}} \mathcal{A}(x) \ , \quad f(x) = \frac{1}{\mathcal{A}_{\min}} F(x) \ . \tag{2.1.4}$$

Clearly then, α is piecewise constant with respect to the partition $\{\Omega_i : i = 0, \ldots, m\}$ and $\alpha(x) \geq 1$ for all $x \in \Omega$.

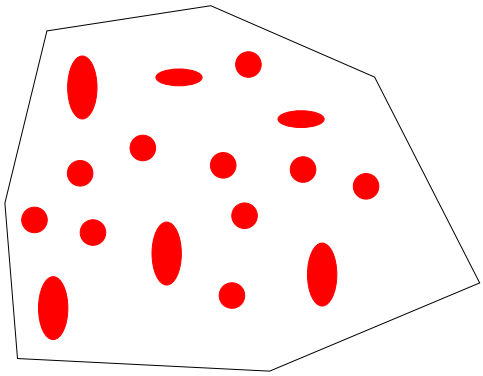


Figure 2.1: A domain with many inclusions.

Letting α_i denote the restriction of α to Ω_i, our analysis focuses on the proof of a robust optimal order of convergence in each of these two "high contrast" cases

characterized by a large contrast parameter $\widehat{\alpha}$

$$\text{Case I}: \qquad \widehat{\alpha} := \min_{i=1,\ldots,m} \alpha_i \rightarrow \infty \ , \quad \alpha_0 = 1 \tag{2.1.5}$$

$$\text{Case II}: \qquad \widehat{\alpha} := \alpha_0 \rightarrow \infty \ , \quad \max_{i=1,\ldots,m} \alpha_i \ \leq \ K, \tag{2.1.6}$$

for some constant K. In Case I, the inclusions have high permeability compared to the background matrix, while Case II contains the converse configuration. In Section 2.3 and 2.4, we show that in both cases, our method satisfies the estimates

$$|u - u_h^{\text{MS}}|_{H^1(\Omega)} \ \leq \ |u - u_h^{\text{MS}}|_{H^1(\Omega),\alpha} \ \leq \ Ch \left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L^2(\Omega)} \right]^{1/2} \tag{2.1.7}$$

and

$$\|u - u_h^{\text{MS}}\|_{L_2(\Omega)} \ \leq \ Ch^2 \left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L^2(\Omega)} \right]^{1/2} \ , \tag{2.1.8}$$

where the constant C is independent of h and of the contrast parameter $\widehat{\alpha}$.

This chapter is organized as following. In Section 2.2.1, we explain the main idea behind our method and illustrate the analysis in the special case when the coarse mesh can be drawn to enclose each of the inclusions $\Omega_i : i = 1, \ldots, m$ in Section 2.3. The main results of the work are obtained in Section 2.4, where we construct the multiscale basis functions for the case when the coefficient interface intersects a typical triangular coarse mesh element. Here we present a detailed analysis, obtaining the estimates (2.1.7) and (2.1.8) under certain technical assumptions. Numerical experiments which illustrate these results are provided in Section 2.5. Two new technical results which are crucial to the analysis are included as appendices: the first is a scaled version of the trace theorem on a triangular annulus, while the second is a regularity theorem (due to N. Babych, I. Kamotski and V.P. Smyshlyaev of the University of Bath) for the exact solution of (2.1.2) in the high-contrast case. These should be of independent interest.

2.2 The Key Idea and Regularity Theorem

2.2.1 The Key Idea

For any measurable subset of $D \subset \Omega$, define the local version of a:

$$a_D(v,w) = \int_D \alpha \nabla v \cdot \nabla w \ .$$

For a suitable index set $\mathscr{I}_h(D)$, let $\mathcal{N}_h(D) = \{x_p : p \in \mathscr{I}_h(D)\}$ denote the nodes of the mesh $\mathcal{T}_h$ which lie in D. For any triangular element $\tau \in \mathcal{T}_h$ (assumed to contain its boundary), $\mathcal{N}_h(\tau) = \{x_p : p \in \mathscr{I}_h(\tau)\}$ is the set containing the three nodes of τ. For each $p \in \mathscr{I}_h(\tau)$ we construct nodal basis functions Φ_p^{MS} whose restriction $\Phi_{p,\tau}^{\mathrm{MS}}$ to each $\tau \in \mathcal{T}_h$ is required to solve the "*local problem*":

$$a_\tau(\Phi_{p,\tau}^{\mathrm{MS}}, v) = 0 \ , \quad \text{for all} \quad v \in H_0^1(\tau) \ , \tag{2.2.1}$$

together with a suitable boundary condition:

$$\Phi_{p,\tau}^{\mathrm{MS}} = \phi_{p,\tau}, \quad \text{on} \quad \partial\tau \ , \quad \text{with} \quad \phi_{p,\tau}(x_q) = \delta_{p,q}, \quad \text{for all} \quad p,q \in \mathscr{I}_h(\tau) \tag{2.2.2}$$

where $\phi_{p,\tau} \in \mathcal{C}(\partial\tau)$ and

$$\sum_{p \in \mathscr{I}_h(\tau)} \phi_{p,\tau} \ = \ 1 \quad \text{on} \quad \partial\tau \ . \tag{2.2.3}$$

In general for each $p \in \mathscr{I}_h(\overline{\Omega})$ the boundary data in (2.2.2) has to be prescribed and the local problems (2.2.1) may have to be solved (e.g. on a subgrid). We will see that there is a bounded number of these local problems for each p, independent of the coarse mesh diameter. However if α is constant on τ then the boundary condition is chosen so that $\Phi_{p,\tau}^{\mathrm{MS}}$ is simply the linear hat function on τ centred at x_p. Observe that under conditions (2.2.1), (2.2.2) and (2.2.3),

$$\sum_{p \in \mathscr{I}_h(\tau)} \Phi_{p,\tau}^{\mathrm{MS}} \ = \ 1 \qquad \text{on} \quad \tau \ . \tag{2.2.4}$$

We extend the basis functions Φ_p^{MS} trivially on Ω and define the multiscale finite element space $\mathcal{V}_h^{\mathrm{MS}} = \mathrm{span}\{\Phi_p^{\mathrm{MS}}\}$. From the basis functions we construct the nodal

interpolation operator:

$$\mathcal{I}_h^{\mathrm{MS}} v = \sum_{p \in \mathscr{I}_h(\overline{\Omega})} v(x_p) \Phi_p^{\mathrm{MS}} \ ,$$

defined for all $v \in C(\overline{\Omega})$. Note that from (2.2.4) it follows that

$$\mathcal{I}_h^{\mathrm{MS}} 1 = 1 \qquad \text{on} \quad \overline{\Omega} \ . \tag{2.2.5}$$

The multiscale finite element solution u_h^{MS} satisfies $a(u_h^{\mathrm{MS}}, v_h^{\mathrm{MS}}) = (f, v_h^{\mathrm{MS}})_{L^2(\Omega)}$, for all $v_h^{\mathrm{MS}} \in \mathcal{V}_h^{\mathrm{MS}}$, which gives us the usual optimality estimate

$$|u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \ \leq \ |u - \mathcal{I}_h^{\mathrm{MS}} u|_{H^1(\Omega),\alpha} \ . \tag{2.2.6}$$

To estimate the right-hand side of (2.2.6), we note that, since the basis functions satisfy the homogeneous equation (2.2.1), and since u solves (2.1.2), the local interpolation error $E_h^{\mathrm{MS}} := u - \mathcal{I}_h^{\mathrm{MS}} u$ satisfies, for all $\tau \in \mathcal{T}_h$,

$$a_\tau(E_h^{\mathrm{MS}}, v) = (f, v)_{L^2(\tau)} \ , \quad \text{for all} \quad v \in H_0^1(\tau) \ . \tag{2.2.7}$$

An estimate for $|E_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha}$ suitable for inserting in the right-hand side of (2.2.6) can be obtained from the following lemma.

Lemma 2.2.1. *Suppose D is a Lipschitz subdomain of Ω and suppose that $\phi \in H^1(D)$ satisfies*

$$a_D(\phi, v) \ = \ (f, v)_{L_2(D)} \quad \textit{for all} \quad v \in H_0^1(D) \ . \tag{2.2.8}$$

Then for any $\tilde{\phi} \in H^1(D)$ such that the trace of $\tilde{\phi} - \phi$ vanishes on ∂D,

$$|\phi|_{H_1(D),\alpha} \ \leq \ |\tilde{\phi}|_{H^1(D),\alpha} \ + \ C \ \mathrm{diam}(D) \|f\|_{L_2(D)} \ ,$$

where C is independent of $\phi, \tilde{\phi}$, the diameter of D and $\widehat{\alpha}$.

Proof. Let ϕ^* be the unique solution of the problem

$$a_D(\phi^*, v) = 0 \quad \text{for all} \quad v \in H_0^1(D) \ , \tag{2.2.9}$$

such that the trace of $\phi^* - \phi$ vanishes on ∂D. Then $\phi - \phi^* \in H_0^1(D)$ and, subtracting

(2.2.9) from (2.2.8),

$$a_D(\phi - \phi^*, v) = (f, v)_{L_2(D)} \quad \text{for all} \quad v \in H_0^1(D) \ .$$

Then

$$\begin{aligned} |\phi - \phi^*|^2_{H^1(D),\alpha} &= a_D(\phi - \phi^*, \phi - \phi^*) \ = \ a_D(\phi, \phi - \phi^*) \ = \ (f, \phi - \phi^*)_{L_2(D)} \\ &\leq \ \|f\|_{L_2(D)}\|\phi - \phi^*\|_{L_2(D)} \leq C \ \mathrm{diam}(D)\|f\|_{L_2(D)}|\phi - \phi^*|_{H^1(D),\alpha} \ , \end{aligned}$$

where in the last step we used the Poincaré-Friedrichs inequality and the fact $\widehat{\alpha} \geq 1$. Hence

$$|\phi - \phi^*|_{H^1(D),\alpha} \leq C \ \mathrm{diam}(D)\|f\|_{L_2(D)},$$

and

$$|\phi|_{H^1(D),\alpha} \ \leq \ |\phi^*|_{H^1(D),\alpha} \ + \ C \ \mathrm{diam}(D)\|f\|_{L_2(D)}.$$

On the other hand, (2.2.9) implies the minimality of the energy norm of ϕ^*, i.e $|\phi^*|_{H^1(D),\alpha} \leq |\tilde{\phi}|_{H^1(D),\alpha}$ for all $\tilde{\phi}$ satisfying the same boundary conditions as ϕ and the result follows. □

Recalling (2.2.7), and using Lemma 2.2.1 to bound the right-hand side of (2.2.6), we obtain

Theorem 2.2.2.

$$|E_h^{\mathrm{MS}}|_{H^1(\tau),\alpha} \ \leq \ |\widetilde{E}_h^{\mathrm{MS}}|_{H^1(\tau),\alpha} \ + \ Ch_\tau\|f\|_{L_2(\tau)} \ , \tag{2.2.10}$$

and

$$|u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \ \leq \ C\Big[\sum_{\tau \in \mathcal{T}^H} \Big(|\widetilde{E}_h^{\mathrm{MS}}|^2_{H^1(\tau),\alpha} \ + \ h_\tau^2\|f\|^2_{L_2(\tau)} \Big) \Big]^{1/2} \ , \tag{2.2.11}$$

where $\widetilde{E}_h^{\mathrm{MS}}$ *is* any *function whose trace coincides with the trace of* E_h^{MS} *on* $\partial\tau$ *and* C *is a generic constant independent of* $\mathcal{T}_h$, f, u *and* α.

Note that, although simple, Theorem 2.2.2 represents a genuine extension of standard estimates. For example, if α is constant on each $\tau \in \mathcal{T}_h$, then the multiscale basis functions coincide with linear basis functions. Setting $\widetilde{E}_h^{\mathrm{MS}} = E_h^{\mathrm{MS}}$ (which now

equals the error in linear interpolation) leads to

$$|u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \;\leq\; Ch\Big[\sum_{\tau\in\mathcal{T}_h}\alpha_\tau|u|^2_{H^2(\tau)} \;+\; \|f\|^2_{L_2(\Omega)}\Big]^{1/2}, \tag{2.2.12}$$

which yields the usual $\mathcal{O}(h)$ estimate in the energy norm with a coefficient dependent asymptotic constant.

However this also demonstrates the possibility that when α_τ is large, small $|u|^2_{H^2(\tau)}$ could provide better estimates with respect to α. This motivates us to study the regularity of the exact solution. Our regularity theory described in the next subsection tells us that this is exactly what does happen. However a deeper use of Theorem 2.2.2 may be envisaged when α varies within an element. Then it turns out to be possible (although not trivial) to define the boundary condition (2.2.2) in such a way that $\sum_{\tau\in\mathcal{T}_h}|\widetilde{E}_h^{\mathrm{MS}}|^2_{H^1(\tau),\alpha} \leq Ch^2$, with constant C independent of α. We explain how this comes about in the context of a special case in Section 2.3. Then in Section 2.4 we extend to the more difficult case where the interface cuts through a mesh element.

From now on we use the following notational conventions.

Notation 2.2.3. We write $g_1 \lesssim g_2$ when there exists a constant C that is independent of u, f, h, α such that $g_1 \leq Cg_2$. Similarly, $g_1 \sim g_2$ means $g_1 \lesssim g_2$ and $g_2 \lesssim g_1$.

Notation 2.2.4. For any suitably smooth function ϕ defined on $\tau \in \mathcal{T}_H$, and any edge e of τ, we define $D_e^s\phi$, for $s \geq 1$ to be the derivative of ϕ of order s along e.

2.2.2 Regularity Estimates for High-Contrast Interface Problems

In this subsection, we discuss the regularity estimate for the interface problem. The regularity theorem (Theorem 2.2.5) indicates that in highly conductive regions (where α is large), the Sobolev norm of the solution behaves like $\mathcal{O}(1/\alpha)$. This observation plays an important role in designing boundary condition for our method and is a key to the analysis. It also substantiates the results of the numerical examples studied in [68, 49, 93].

When the interfaces are smooth enough, the solution of the interface problem is also very smooth in individual regions. However, the global regularity can be very low.[3] More precisely, we have $u|_{\Omega_i} \in H^2(\Omega_i)$ but have only $u|_{\Omega_i} \in H^1(\Omega)$ even when

[3] See Littman et al. [70] and Kellogg [59, 60].

interfaces are sufficiently smooth. And in this case, the solution u admits a priori estimate:

$$\sum_{i=0}^{m} \|u\|_{H^2(\Omega_i)} \leq C\|f\|_{L^2(\Omega)},$$

where C is independent of u, f but depends strongly and implicitly on the jumps in the coefficients across the interface. While the Sobolev regularity of the solution u to the interface problem (2.1.2) is classical, there are relatively few published results which give estimates of how the Sobolev norms of u depend on the contrast parameter $\widehat{\alpha}$. An exception is Huang and Zou [53] which gives a partial result in this direction which we generalize in the following theorem. The proof of the theorem below was proposed to us by N. Babych, I.V. Kamotski and V.P. Smyshlyaev of the University of Bath, UK.

Theorem 2.2.5. *(Babych, Kamotski, Smyshlyaev) Let Ω be either a smooth $\mathcal{C}^\infty$ bounded domain in $\mathbb{R}^2$ or a bounded convex polygon, containing inclusions Ω_i, $i = 1, 2 \ldots, m$, each having a $\mathcal{C}^\infty$ boundary, and define $\Omega_0 = \Omega \setminus \cup_{i=1}^m \overline{\Omega}_i$ as described in Section 2.1. Consider problem* (2.1.2) *and assume that either Case I or Case II (*(2.1.5) *or* (2.1.6)*) holds. Additionally, let $\widetilde{\Gamma}$ denote any closed $\mathcal{C}^\infty$ contour in Ω_0 which encloses all the Ω_i and let $\widetilde{\Omega}_0$ be the domain with boundary $\Gamma \cup \widetilde{\Gamma}$ (see Fig. 2.2 for an illustration in the case $m = 1$). Then we have*

$$|u|_{H^{s+2}(\Omega_i)} \lesssim \frac{1}{\alpha_i} \|f\|_{H^s(\Omega)}, \quad \text{for all} \quad s \geq 0\ ,\ i = 1, 2 \ldots, m\ . \tag{2.2.13}$$

Moreover

$$|u|_{H^2(\Omega_0)} \lesssim \frac{1}{\alpha_0} \|f\|_{L_2(\Omega)}, \tag{2.2.14}$$

and

$$|u|_{H^{2+s}(\widetilde{\Omega}_0)} \lesssim \frac{1}{\alpha_0} \|f\|_{H^s(\Omega)}, \quad \text{for all} \quad s \geq 0\ . \tag{2.2.15}$$

The hidden constants depend on the distance of Γ from $\partial\Omega$.

Proof. We give the proof in Appendix A.2. □

Remark 2.2.6. While the estimates (2.2.14) and (2.2.15) can be recovered from the results in [53], the result (2.2.13) can not be found there, because [53] works only with estimates of norms and proves $\|u\|_{H^{s+2}(\Omega_i)} \lesssim \|f\|_{H^s(\Omega)}$, $i = 1, \ldots, m$. The bounds on

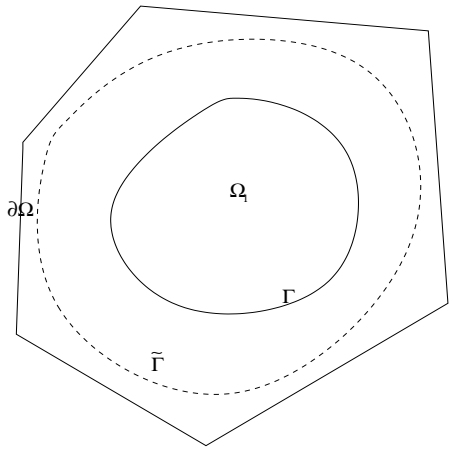


Figure 2.2: A domain Ω in the case $m = 1$. The domain Ω_0 is bounded by Γ and $\partial\Omega$ while $\widetilde{\Omega}_0$ is bounded by Γ and $\tilde{\Gamma}$.

the H^{s+2}-seminorms in (2.2.13) are sharper in Case I and are essential to the analysis in the present work.

2.2.3 Steps of Analysis

The accuracy of our multiscale finite element method depends critically on the construction of accurate local boundary conditions for the multiscale bases independent of the contrast of the coefficients. In the following two sections, we describe how to design the boundary conditions for multiscale basis functions to obtain the optimal convergence rate:

$$|u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \;\leq\; Ch\left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L^2(\Omega)}\right]^{1/2}$$

and

$$\|u - u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \;\leq\; Ch^2\left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L^2(\Omega)}\right]^{1/2} .$$

Before we present the formal analysis, we first explain the main strategies of the analysis. Recall the local interpolation error function $E_h^{\mathrm{MS}} := u - \mathcal{I}_h^{\mathrm{MS}} u$ and the result in Theorem 2.2.2

$$|u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \;\leq\; C\Big[\sum_{\tau\in\mathcal{T}^H}\Big(|\widetilde{E}_h^{\mathrm{MS}}|^2_{H^1(\tau),\alpha} \;+\; h_\tau^2\|f\|^2_{L_2(\tau)}\Big)\Big]^{1/2} ,$$

where $\widetilde{E}_h^{\mathrm{MS}}$ is *any* function whose trace coincides with the trace of E_h^{MS} on $\partial\tau$. The key is to show there exist local boundary conditions and extensions $\widetilde{E}_h^{\mathrm{MS}}$ yielding

$$\sum_{\tau\in\mathcal{T}^H} |\widetilde{E}_h^{\mathrm{MS}}|^2_{H^1(\tau),\alpha} \lesssim h^2 \left(h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L_2(\Omega)} \right).$$

This involves 3 steps:

- Step 1: Construct boundary conditions for basis functions with the desired errors. When inclusions are inside of elements, we use linear functions to define the boundary conditions. Since the error function E_h^{MS} is zero at nodes, we should expect to have $E_h^{\mathrm{MS}} = \mathcal{O}(h_\tau^2)$ along the sides of the element τ under some appropriate regularity condition. See Fig. 2.3 (a).

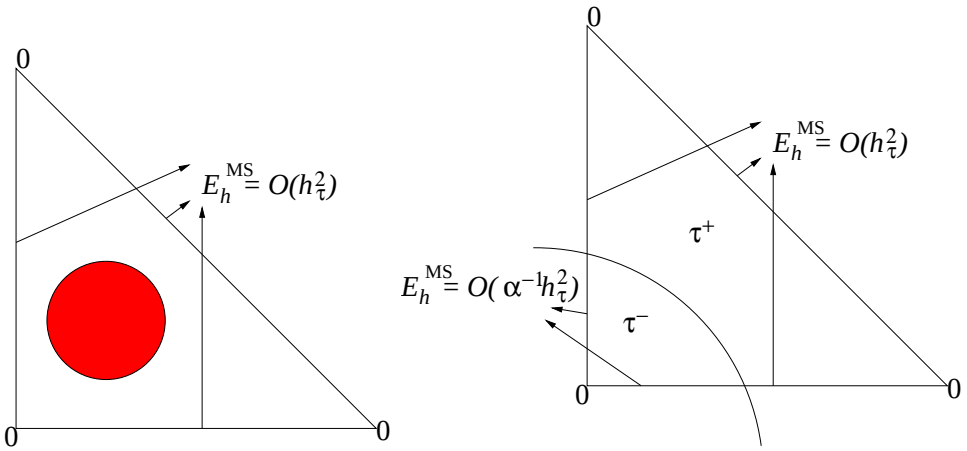
  When an interface cuts through a element, the element is divided into two parts. We denote the part with $\alpha = \widehat{\alpha}$ by τ^- and the other by τ^+. We use piecewise linear functions to construct the boundary conditions. We need to have $E_h^{\mathrm{MS}} = \mathcal{O}(\alpha^{-1}h_\tau^2)$ on $\partial\tau^-\cap\partial\tau$ and $E_h^{\mathrm{MS}} = \mathcal{O}(h_\tau^2)$ on $\partial\tau^+\cap\partial\tau$. The factor, α^{-1}, on $\partial\tau^-\cap\partial\tau$ is important to eliminate the dependence on the contrast of the coefficient. See Fig. 2.3 (b). This step is most crucial for this work. More details are discussed in Section 2.4.

- Step 2: Construct suitable extensions inside elements that maintain the order of the error on the boundary. This can be done by applying the trace theorem. The H^1-semi norm of the extension can be bounded by boundary values which are controlled in Step 1.

- Step 3: Use a sharp regularity theorem (Theorem 2.2.5) to obtain a bound on the error coefficient, which contains $\widehat{\alpha}$ and the Sobolev norm of u, in terms of the Sobolev norm of f.

2.3 Error Analysis for Inclusions Inside Elements

In this section we assume that each of the inclusions Ω_i are enclosed inside elements of $\mathcal{T}_h$. More precisely, for any $\tau \in \mathcal{T}_h$ and any sufficiently small $\varepsilon > 0$, let us define $\tau^\varepsilon = \{x \in \tau : \mathrm{dist}(x, \partial\tau_i) \leq \varepsilon\}$. Then, for each $i = 1, \ldots, m$, we assume that there



(a) An inclusion inside an element. (b) An inclusion cut through an element.

Figure 2.3: Illustrations for Step 1. The error function E_h^{MS} takes 0 at 3 nodes of the element and the boundary conditions are designed such that E_h^{MS} has the desired error bounds on the boundary of the elements.

exist $\tau_i \in \mathcal{T}_h$ and $\varepsilon_i > 0$ such that

$$\Omega_i \subseteq \ \tau_i \backslash \tau_i^{\varepsilon_i} \ . \tag{2.3.1}$$

That is, the boundary of Ω_i lies at least a distance ε_i from the boundary of τ_i. Note that any element could contain more than one inclusion. Our estimates will depend on the following measure of the relative size of ε_i:

$$\delta \ := \ \max_{i=1,\ldots,m} \frac{h_{\tau_i}}{\varepsilon_i} \ . \tag{2.3.2}$$

Note that $\delta \geq 1$. Then we have the following theorem:

Theorem 2.3.1. *Suppose the boundary condition in* (2.2.2) *is linear on the boundary of each element $\tau \in \mathcal{T}_h$. Then, for $f \in H^{1/2}(\Omega)$,*

$$(i) \quad |u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \ \lesssim \delta^3 h \left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L_2(\Omega)} \right]^{1/2}, \tag{2.3.3}$$

$$(ii) \quad \|u - u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \ \lesssim \delta^6 h^2 \left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L_2(\Omega)} \right]^{1/2}. \tag{2.3.4}$$

Proof. We give the proof in Case I (see (2.1.5)). Thus we assume $\alpha_i \geq \widehat{\alpha} \to \infty$ for $i = 1, \ldots, m$ and $\alpha_0 = 1$. Case II is discussed in Remark 2.3.2.

First consider any element $\tau_i \in \mathcal{T}_h$ which contains an inclusion Ω_i. We construct $\widetilde{E}_h^{\mathrm{MS}}$ on τ_i by defining it to be zero on $\tau_i \backslash \tau_i^{\varepsilon_i}$, equal to E_h^{MS} on $\partial\tau_i$ and extending

into $\tau_i^{\varepsilon_i}$ using Lemma A.1.1 of Appendix A (scaled to τ_i). By combining this with standard estimates for linear approximation in 1D, we obtain

$$\begin{aligned} |\widetilde{E}_h^{\rm MS}|^2_{H^1(\tau_i),\alpha} &= |\widetilde{E}_h^{\rm MS}|^2_{H^1(\tau_i^{\varepsilon_i})} \\ &\lesssim \left(\frac{h_{\tau_i}}{\varepsilon_i}\right)^2 h_{\tau_i}^{-1}\|E_h^{\rm MS}\|^2_{L_2(\partial\tau_i)} + \left(\frac{h_{\tau_i}}{\varepsilon_i}\right)^2 h_{\tau_i}|E_h^{\rm MS}|^2_{H^1(\partial\tau_i)} \\ &\lesssim \delta^2 h_{\tau_i}^3 \sum_{e\in\mathcal{E}(\tau)} \|D_e^2 u\|^2_{L_2(e)} , \end{aligned}$$

where $\mathcal{E}(\tau)$ denotes the edges of τ. On the other hand if $\tau \in \mathcal{T}_h$ contains no inclusion, then with $\widetilde{E}_h^{\rm MS} := E_h^{\rm MS}$, it is easy to show that this estimate remains true with δ replaced by 1, thus

$$|\widetilde{E}_h^{\rm MS}|^2_{H^1(\tau),\alpha} \lesssim \delta^2 h_\tau^3 \sum_{e\in\mathcal{E}(\tau)} \|D_e^2 u\|^2_{L_2(e)} , \quad \text{for all} \quad \tau \in \mathcal{T}_h . \tag{2.3.5}$$

Recalling (2.2.11), we now see that (2.3.5) allows us to estimate $|u - u_h^{\rm MS}|_{H^1(\Omega),\alpha}$ in terms of the data f and certain derivatives of the solution u along edges which (in this case) lie entirely in Ω_0. In order to prove robustness to the contrast $\widehat{\alpha}$, we now estimate these edge derivatives in terms of Sobolev norms of u in Ω_0, which we can in turn estimate independently of $\widehat{\alpha}$, using the regularity theory in the Appendix. The required technical argument, which we now give, leads to (2.3.16).

First we recall the trace theorem for polygons (e.g., [47, Theorem 1.5.2.1]), which, after scaling to any element $\tau \in \mathcal{T}_h$, reads

$$|v|^2_{H^1(e)} \lesssim h_\tau^{-3}\|v\|^2_{L_2(\tau)} + h_\tau^{-1}|v|^2_{H^1(\tau)} + |v|^2_{H^{3/2}(\tau)} , \quad \text{for all} \quad v \in H^{3/2}(\tau) .$$

Moreover, replacing v by $v - \gamma$ where γ is an arbitrary constant, and then invoking the Poincaré inequality on τ, we obtain the simpler estimate

$$|v|^2_{H^1(e)} \lesssim h_\tau^{-1}|v|^2_{H^1(\tau)} + |v|^2_{H^{3/2}(\tau)} \quad \text{for all} \quad v \in H^{3/2}(\tau) \quad \text{and all} \quad \tau \in \mathcal{T}_h . \tag{2.3.6}$$

Now return to the case where $\tau = \tau_i$, an element which contains an inclusion Ω_i. Choose ϕ to be a $\mathcal{C}^\infty$ cut-off function which vanishes on $\tau_i \backslash \tau_i^{\varepsilon_i}$, has the value 1 on $\partial\tau_i$ and satisfies $\|D^\beta\phi\|_{L^\infty(\tau_i)} \lesssim \varepsilon_i^{-|\beta|}$ for all multindices β. Then choose any edge e of τ_i

and any constant γ. Using the fact that $\phi\gamma$ is constant on e and (2.3.6), we obtain

$$\|D_e^2 u\|^2_{L_2(e)} \;=\; |\phi(D_e u-\gamma)|^2_{H^1(e)} \;\lesssim\; h_{\tau_i}^{-1}|\phi(D_e u-\gamma)|^2_{H^1(\tau_i)} \;+\; |\phi(D_e u-\gamma)|^2_{H^{3/2}(\tau_i)}\,. \tag{2.3.7}$$

To estimate the right-hand side, we recall the Poincaré inequality for an annulus (Lemma A.1.2) and scale it to τ_i to obtain a constant $\gamma\in\mathbb{R}$ such that

$$\|v-\gamma\|_{L_2(\tau_i^{\varepsilon_i})} \;\lesssim\; h_{\tau_i}|v|_{H^1(\tau_i^{\varepsilon_i})}\quad,\quad \text{for all}\quad v\in H^1(\tau_i^{\varepsilon_i})\,, \tag{2.3.8}$$

with a hidden constant that is independent of both h_{τ_i} and ε_i .

Now, to estimate the first term on the right-hand side of (2.3.7), use the above estimates for the derivatives of ϕ and (2.3.8), to obtain

$$\begin{aligned} h_{\tau_i}^{-1}|\phi(D_e u-\gamma)|^2_{H^1(\tau_i)} \quad &\lesssim \quad h_{\tau_i}^{-1}\varepsilon_{\tau_i}^{-2}\|D_e u-\gamma\|^2_{L_2(\tau_i^{\varepsilon_i})} \;+\; h_{\tau_i}^{-1}|u|^2_{H^2(\tau_i^{\varepsilon_i})} \\ &\lesssim \quad h_{\tau_i}^{-1}\delta^2|u|^2_{H^2(\tau_i^{\varepsilon_i})}\,. \end{aligned} \tag{2.3.9}$$

The estimation of the second term on the right-hand side of (2.3.7) is slightly more involved. For any multi-index β of order 1, we have

$$|\phi(D_e u-\gamma)|^2_{H^{3/2}(\tau_i)} \;\lesssim\; \sum_{|\beta|=1}|(D^\beta\phi)(D_e u-\gamma)|^2_{H^{1/2}(\tau_i^{\varepsilon_i})} \;+\; \sum_{|\beta|=1}|\phi D^\beta D_e u|^2_{H^{1/2}(\tau_i^{\varepsilon_i})}\,. \tag{2.3.10}$$

Now for all $\psi\in\mathcal{C}^\infty(\tau_i^{\varepsilon_i})$ and $v\in H^{1/2}(\tau_i^{\varepsilon_i})$, elementary arguments show:

$$|\psi v|^2_{H^{1/2}(\tau_i^{\varepsilon_i})} \;\le\; \|\psi\|^2_{L_\infty(\tau_i^{\varepsilon_i})}|v|^2_{H^{1/2}(\tau_i^{\varepsilon_i})} \;+\; h_{\tau_i}\|\nabla\psi\|^2_{L_\infty(\tau_i^{\varepsilon_i})}\|v\|^2_{L_2(\tau_i^{\varepsilon_i})}\,. \tag{2.3.11}$$

Hence, the first term on the right-hand side of (2.3.10) is estimated by

$$|(D^\beta\phi)(D_e u-\gamma)|^2_{H^{1/2}(\tau_i^{\varepsilon_i})} \;\lesssim\; \varepsilon_{\tau_i}^{-2}|D_e u-\gamma|^2_{H^{1/2}(\tau_i^{\varepsilon_i})} + \varepsilon_{\tau_i}^{-4}h_{\tau_i}\|D_e u-\gamma\|^2_{L_2(\tau_i^{\varepsilon_i})}\,. \tag{2.3.12}$$

Interpolating (2.3.8) with the corresponding estimate for $\|D_e u-\gamma\|_{H^1(\tau_i^{\varepsilon_i})}$ we obtain

$$|D_e u-\gamma|_{H^{1/2}(\tau_i^{\varepsilon_i})} \;\lesssim\; h_{\tau_i}^{1/2}|D_e u|_{H^1(\tau_i^{\varepsilon_i})}\,. \tag{2.3.13}$$

Combining (2.3.8) and (2.3.13) with (2.3.12), we have

$$|(D^\beta\phi)(D_e u-\gamma)|^2_{H^{1/2}(\tau_i^{\varepsilon_i})} \lesssim\; (\varepsilon_{\tau_i}^{-2}h_{\tau_i}+\varepsilon_{\tau_i}^{-4}h_{\tau_i}^3)|u|^2_{H^2(\tau_i^{\varepsilon_i})} \lesssim\; h_{\tau_i}^{-1}\delta^4|u|^2_{H^2(\tau_i^{\varepsilon_i})}\,. \tag{2.3.14}$$

The second term on the right-hand side of (2.3.10) is also estimated by (2.3.11):

$$|\phi D^\beta D_e u|^2_{H^{1/2}(\tau_i^{\varepsilon_i})} \lesssim |u|^2_{H^{5/2}(\tau_i^{\varepsilon_i})} + h_{\tau_i}\varepsilon_{\tau_i}^{-2}|u|^2_{H^2(\tau_i^{\varepsilon_i})} \lesssim |u|^2_{H^{5/2}(\tau_i^{\varepsilon_i})} + h_{\tau_i}^{-1}\delta^2|u|^2_{H^2(\tau_i^{\varepsilon_i})} . \tag{2.3.15}$$

Then combining (2.3.14) and (2.3.15) with (2.3.10) we have

$$|\phi(D_e u - \gamma)|^2_{H^{3/2}(\tau_i)} \lesssim h_{\tau_i}^{-1}\delta^4|u|^2_{H^2(\tau_i^{\varepsilon_i})} + |u|^2_{H^{5/2}(\tau_i^{\varepsilon_i})} .$$

Combining this with (2.3.9) and (2.3.7), we have

$$\|D_e^2 u\|^2_{L_2(e)} \lesssim h_{\tau_i}^{-1}\delta^4|u|^2_{H^2(\tau_i^{\varepsilon_i})} + |u|^2_{H^{5/2}(\tau_i^{\varepsilon_i})} .$$

By a direct application of (2.3.6), this estimate also holds (in fact with δ replaced by 1) when τ does not contain an inclusion , so that

$$\|D_e^2 u\|^2_{L_2(e)} \lesssim h_\tau^{-1}\delta^4|u|^2_{H^2(\tau\cap\Omega_0)} + |u|^2_{H^{5/2}(\tau\cap\Omega_0)} , \quad \text{for all} \quad e \in \mathcal{E}(\tau) \quad \text{and all} \quad \tau \in \mathcal{T}_h .$$

Combining this with (2.3.5) yields, for all $\tau \in \mathcal{T}_h$,

$$|\widetilde{E}_h^{\mathrm{MS}}|^2_{H^1(\tau),\alpha} \lesssim \delta^2 h_\tau^2 \left[\delta^4|u|^2_{H^2(\tau\cap\Omega_0)} + h_\tau|u|^2_{H^{5/2}(\tau\cap\Omega_0)}\right]. \tag{2.3.16}$$

Combining with (2.2.11) and employing the regularity theory from the Appendix yields the result (i).

Let $w \in H_0^1(\Omega)$ be the solution of

$$a(w, v) = (u - u_h^{\mathrm{MS}}, v)_{L_2(\Omega)} \quad \text{for all} \quad v \in H_0^1(\Omega) \tag{2.3.17}$$

and let $w_h^{\mathrm{MS}} \in \mathcal{V}_h^{\mathrm{MS}}$ satisfy:

$$a(w_h^{\mathrm{MS}}, v_h^{\mathrm{MS}}) = (u - u_h^{\mathrm{MS}}, v_h^{\mathrm{MS}})_{L_2(\Omega)} \quad \text{for all} \quad v_h^{\mathrm{MS}} \in \mathcal{V}_h^{\mathrm{MS}} . \tag{2.3.18}$$

Applying (2.3.3), then the interpolation theorem for $H^{1/2}(\Omega)$ and finally the arithmetic-

geometric mean inequality ($ab \le a^2/2 + b^2/2$) yields

$$\begin{aligned}
|w - w_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} &\lesssim \delta^3 h \left[h|u - u_h^{\mathrm{MS}}|^2_{H^{1/2}(\Omega)} \ + \ \|u - u_h^{\mathrm{MS}}\|^2_{L_2(\Omega)} \right]^{1/2} \\
&\lesssim \delta^3 h \left[h|u - u_h^{\mathrm{MS}}|_{H^1(\Omega)} \|u - u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \ + \ \|u - u_h^{\mathrm{MS}}\|^2_{L_2(\Omega)} \right]^{1/2} \\
&\lesssim \delta^3 h \left[h^2|u - u_h^{\mathrm{MS}}|^2_{H^1(\Omega)} \ + \ \|u - u_h^{\mathrm{MS}}\|^2_{L_2(\Omega)} \right]^{1/2} \\
&\lesssim \delta^3 h \left[h|u - u_h^{\mathrm{MS}}|_{H^1(\Omega)} \ + \ \|u - u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \right] . \qquad (2.3.19)
\end{aligned}$$

Hence , taking $v = u - u_h^{\mathrm{MS}}$ in (2.3.17), we get

$$\begin{aligned}
\|u - u_h^{\mathrm{MS}}\|^2_{L_2(\Omega)} &= a(w, u - u_h^{\mathrm{MS}}) \ = \ a(w - w_h^{\mathrm{MS}}, u - u_h^{\mathrm{MS}}) \\
&\le |w - w_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} |u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \ ,
\end{aligned}$$

and combining this with (2.3.19) (and recalling $\alpha \ge 1$), we have

$$\|u - u_h^{\mathrm{MS}}\|^2_{L_2(\Omega)} \ \le C\delta^3 h \left[h|u - u_h^{\mathrm{MS}}|^2_{H^1(\Omega),\alpha} \ + \ |u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \|u - u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \right] \qquad (2.3.20)$$

for some constant $C > 0$. Now by the arithmetic-geometric mean inequality again we have

$$C\delta^3 h |u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \|u - u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \le \ \frac{1}{2} C^2 \delta^6 h^2 |u - u_h^{\mathrm{MS}}|^2_{H^1(\Omega),\alpha} \ + \frac{1}{2} \|u - u_h^{\mathrm{MS}}\|^2_{L_2(\Omega)}$$

and substitution into (2.3.20) yields

$$\begin{aligned}
\frac{1}{2}\|u - u_h^{\mathrm{MS}}\|^2_{L_2(\Omega)} &\le C\delta^3 h^2 |u - u_h^{\mathrm{MS}}|^2_{H^1(\Omega),\alpha} \ + \ \frac{1}{2} C^2 \delta^6 h^2 |u - u_h^{\mathrm{MS}}|^2_{H^1(\Omega),\alpha} \\
&= C\delta^3 (1 + (C/2)\delta^3) h^2 |u - u_h^{\mathrm{MS}}|^2_{H^1(\Omega),\alpha} \ ,
\end{aligned}$$

which combined with (2.3.3) leads to the desired result (2.3.4). □

Remark 2.3.2. For Case II, we can use the same idea to prove

$$(i) \quad |u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \ \lesssim \delta^3 h \left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L_2(\Omega)} \right]^{1/2}, \qquad (2.3.21)$$

$$(ii) \quad \|u - u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \ \lesssim \delta^6 h^2 \left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L_2(\Omega)} \right]^{1/2}. \qquad (2.3.22)$$

We construct $\widetilde{E}_h^{\mathrm{MS}}$ as in the proof of Theorem 2.3.1, and notice that $|\widetilde{E}_h^{\mathrm{MS}}|_{H^1(\tau_i),\alpha} = \widehat{\alpha}^{1/2} |\widetilde{E}_h^{\mathrm{MS}}|_{H^1(\tau_i)}$. The regularity result in Theorem 2.2.5 leads to (2.3.21). The same

duality argument gives (2.3.22).

The estimates above can be extended to the inhomogeneous Dirichlet case. This is discussed in detail in a more general context in Remark 2.4.19.

2.4 Error Analysis for Elements Which Intersect Inclusions

Now we consider the case in which the interface may intersect with the boundaries of some of the mesh elements. Recalling Theorem 2.2.2, we have to show that for each element τ there is a boundary condition for the multiscale basis functions on $\partial\tau$ such that $E_h^{\mathrm{MS}} := u - \mathcal{I}_h^{\mathrm{MS}} u$ has an extension $\widetilde{E}_h^{\mathrm{MS}}$ from $\partial\tau$ into τ with a suitably bounded energy.

In Section 2.4.1 we present some qualitative properties of the exact solution u of (2.1.2) in the generic case when the interface intersects two edges of τ. These relate various derivatives of u at the points of edge-interface intersection, plus controllable remainders. These relations motivate the interior boundary conditions for the multiscale basis functions presented in Section 2.4.2. In particular it is explained how the boundary conditions can be found by solving a 6×6 linear system for each element which intersects the interface. In Theorem 2.4.9 we estimate E_h^{MS} on each element boundary.

The interior error is considered in Section 2.4.3. First, Lemma 2.4.15 uses the result of Theorem 2.4.9 to prove the existence of an extension $\widetilde{E}_h^{\mathrm{MS}}$ of E_h^{MS} with suitably bounded energy; this leads to Theorem 2.4.16, which proves a suitable error estimate for the energy norm of E_h^{MS} on each element τ, by using Theorem 2.2.2 and the estimate for the extension proved in Theorem 2.4.9.

The procedure which we describe constructs interior boundary conditions element by element and does not naturally lead to conforming elements. However conformity can be regained by local averaging. This is described in Section 2.4.4, where the main theorem, Theorem 2.4.18 is proved. An important observation, discussed in Section 2.4.2, is that in the case where the interface intersects the element edges orthogonally, our boundary condition coincides with the "oscillatory boundary conditions" proposed in [51].

2.4.1 Properties of the Exact Solution

In this subsection we derive some properties of the exact solution of (2.1.2) in an element τ through which an intersection cuts. To avoid some technicalities in the theory we shall make the following geometrical assumption on elements which intersect the interface.

Assumption 2.4.1. When the interface Γ intersects an element τ we shall assume that it subdivides τ into two parts. We label the vertices of τ as x_1, x_2, x_3 in such a way that Γ intersects $\partial\tau$ at points which we denote by y_i in the edges $\overline{x_3x_i}$, for $i = 1, 2$. We let β denote the angle of τ subtended at x_3 and we let τ^- denote the part where $\alpha \geq \widehat{\alpha}$ and τ^+ denote the part where $\alpha \leq K$ (see (2.1.5), (2.1.6)). A typical situation is depicted in Figure 2.4. Letting r_i^- and r_i^+ denote, respectively, the length of the line segments $e_i \cap \tau^-$ and $e_i \cap \tau^+$, we shall assume that there exist positive constants $0 \leq \underline{R} < \overline{R} \leq 1$ and $0 < B < \pi$ such that, for all meshes (characterised by mesh parameter h),

$$\underline{R}h_\tau \leq \min\{r_i^-, r_i^+\} \leq \max\{r_i^-, r_i^+\} \leq \overline{R}h_\tau \quad \text{for} \quad i = 1, 2 \quad \text{and} \quad \pi - B \geq \beta \geq B\,. \tag{2.4.1}$$

For $i = 1, 2$ we define $\theta_i \in (-\pi/2, \pi/2)$ to be the unique angle such that

$$\mathbf{e}_i = \cos\theta_i\ \mathbf{n}_i + \sin\theta_i\ \mathbf{t}_i\ . \tag{2.4.2}$$

We also assume that Γ is not tangential to either of the edges e_i, i.e.

$$|\theta_i| \leq \pi/2 - T \quad \text{for some} \quad T > 0\ , \tag{2.4.3}$$

The hidden constants in the estimates below may depend on $\beta, \underline{R}, \overline{R}$, B and T.

In many cases where the element intersects the interface in a different way, we can always find a refinement to reduce to cases satisfying Assumption 2.4.1 , see, e.g. Figure 2.5.

Referring again to Figure 2.4, for $i = 1, 2$, we let $\mathbf{e}_i$ denote the unit vector directed from x_3 to x_i and let $\mathbf{n}_i$ and $\mathbf{t}_i$ denote, respectively, the unit normal and the unit tangent to Γ at y_i. These are uniquely determined by requiring that $\mathbf{n}_i$ is directed

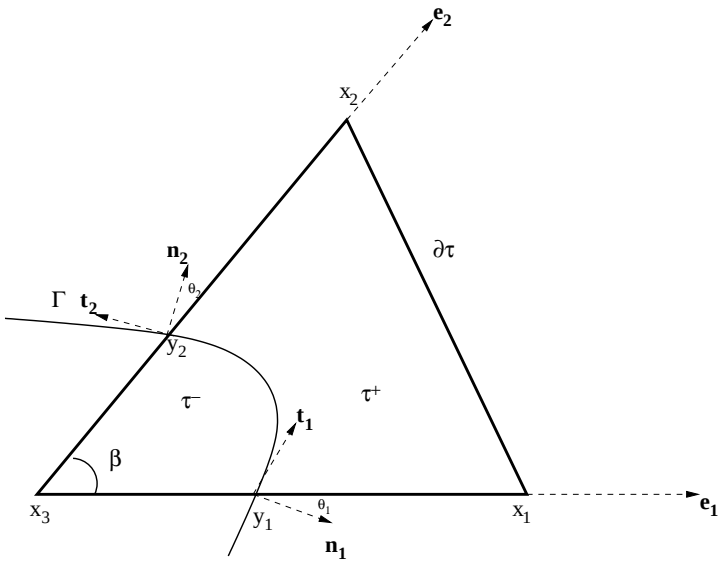


Figure 2.4: An interface cutting through an element: Here τ^- denotes the part where the coefficient α is large.

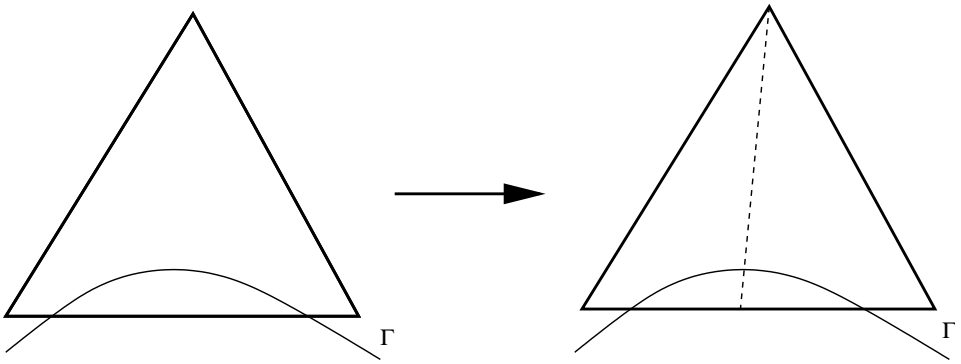


Figure 2.5: After a simple refinement, we can reduce the element to the case we consider.

outward from τ^- and that $\mathbf{t}_i = R_{\pi/2}\mathbf{n}_i$, where R_ϕ is the rotation matrix

$$R_\phi = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} .$$

Note also that $\mathbf{e}_i = R_{\theta_i}\ \mathbf{n}_i$. In particular $\mathbf{e}_i = \mathbf{n}_i$ if and only if $\theta_i = 0$.

Later we use the easily derived relations:

$$R_\phi\ \mathbf{n}_i = \cos\phi\ \mathbf{n}_i + \sin\phi\ \mathbf{t}_i \quad \text{and} \quad R_\phi\ \mathbf{t}_i = -\sin\phi\ \mathbf{n}_i + \cos\phi\ \mathbf{t}_i \ , \qquad (2.4.4)$$

for $i = 1, 2$ and any ϕ. From these it follows that $\mathbf{n}_2 = R_{-\theta_2}\ \mathbf{e}_2 = R_{-\theta_2+\beta}\ \mathbf{e}_1 = R_{\theta_1-\theta_2+\beta}\ \mathbf{n}_1$, and, similarly, $\mathbf{t}_2 = R_{\theta_1-\theta_2+\beta}\ \mathbf{t}_1$. Combining these last two relations

with (2.4.4) yields:

$$\mathbf{n}_2 \;=\; \cos(\theta_2-\theta_1-\beta)\;\mathbf{n}_1 \;-\; \sin(\theta_2-\theta_1-\beta)\;\mathbf{t}_1, \tag{2.4.5}$$

and

$$\mathbf{t}_2 \;=\; \sin(\theta_2-\theta_1-\beta)\;\mathbf{n}_1 \;+\; \cos(\theta_2-\theta_1-\beta)\;\mathbf{t}_1. \tag{2.4.6}$$

As explained above, we now study how the solution u of (2.1.2) behaves on $\partial\tau$. This information will be used to construct suitable internal boundary conditions for multiscale basis functions in Section 2.4.2.

Throughout, we denote the restriction of u to $\tau^{\pm}$ by $u^{\pm}$. Also, for any unit vector $\mathbf{v}$ we let $D_v u$ denote the derivative of u in the direction $\mathbf{v}$. The boundary conditions derived in the following section will be motivated by the relationships between the quantities $(D_{e_i}u^{\pm})(y_i)$, for $i=1,2$, where u is the exact solution of (2.1.2). Our first relationship is the following:

Lemma 2.4.2. *Let u be the exact solution of* (2.1.2) *and define the vector $\boldsymbol{\varepsilon} = (\varepsilon_1,\varepsilon_2)^T \in \mathbb{R}^2$ by requiring*

$$r_i^-(D_{e_i}u^-)(y_i) + r_i^+(D_{e_i}u^+)(y_i) \;=\; u(x_i)-u(x_3)+\varepsilon_i\ , \quad i=1,2\ . \tag{2.4.7}$$

Then

$$|\varepsilon_i| \;\lesssim\; h_\tau^{3/2}\;\left(\|D_{e_i}^2u^-\|_{L_2(e_i\cap\tau^-)} + \|D_{e_i}^2u^+\|_{L_2(e_i\cap\tau^+)}\right), \quad \textit{for} \quad i=1,2. \tag{2.4.8}$$

Proof. This follows from straightforward Taylor expansions at the point y_i and the interface matching condition $u^+(y_i)=u^-(y_i)$ plus the fact that $u^{\pm}\in H^2(\tau^{\pm})$. □

Now, when u is known at the three node points x_i, x_2, x_3, then, setting $\varepsilon_i = 0$ in (2.4.7) gives us two equations for (approximations of) the four unknown quantities $(D_{e_i}u^{\pm})(y_i)$, $i=1,2$.

To determine additional equations for these quantities, we use the interface jump conditions for u at y_1, y_2 to obtain in a straightforward way:

Lemma 2.4.3. *Let u be the exact solution of* (2.1.2). *For $i = 1, 2$,*

$$\begin{bmatrix} D_{e_i}u^-(y_i) \\ \\ D_{e_i}u^+(y_i) \end{bmatrix} \;=\; A_{\widehat{\alpha},\theta_i} \begin{bmatrix} D_{n_i}u^-(y_i) \\ \\ D_{t_i}u^-(y_i) \end{bmatrix} \tag{2.4.9}$$

where

$$A_{\widehat{\alpha},\,\theta} \;=\; \begin{bmatrix} \cos\theta & \sin\theta \\ \widehat{\alpha}\cos\theta & \sin\theta \end{bmatrix} . \tag{2.4.10}$$

Proof. The proof is obtained by simply combining the interface conditions:

$$(D_{n_i}u^+)(y_i) \;=\; \widehat{\alpha}(D_{n_i}u^-)(y_i) \quad \text{and} \quad (D_{t_i}u^+)(y_i) \;=\; (D_{t_i}u^-)(y_i), \quad \text{for } i = 1, 2. \tag{2.4.11}$$

with (2.4.2).

□

We can use (2.4.9) to define a relation between $(D_{e_i}u^{\pm})(y_i)$, $i = 1, 2$, provided we have a relation between $(D_{n_1}u^-)(y_1)$, $(D_{t_1}u^-)(y_1)$, $(D_{n_2}u^-)(y_2)$, and $(D_{t_2}u^-)(y_2)$. Such a relation is provided by the following lemma.

Lemma 2.4.4. *Let u be the exact solution of* (2.1.2).

$$\begin{bmatrix} D_{n_2}u^-(y_2) \\ \\ D_{t_2}u^-(y_2) \end{bmatrix} \;=\; R_{\theta_2-\theta_1-\beta} \begin{bmatrix} D_{n_1}u^-(y_1) \\ \\ D_{t_1}u^-(y_1) \end{bmatrix} \;+\; \boldsymbol{\varepsilon}' \;, \tag{2.4.12}$$

where

$$\|\boldsymbol{\varepsilon}'\|_\infty \;\lesssim\; h_\tau^{1/2} \left[\left\| D_{e_2}D_{n_1}u^- \right\|^2_{L_2(e_2\cap\tau^-)} \;+\; \left\| D_{e_1}D_{n_1}u^- \right\|^2_{L_2(e_1\cap\tau^-)} \right]^{1/2} .$$

Proof. From equations (2.4.5) and (2.4.6) we have, for all $x \in \tau^-$,

$$\begin{bmatrix} D_{n_2}u^-(x) \\ \\ D_{t_2}u^-(x) \end{bmatrix} \;=\; R_{\theta_2-\theta_1-\beta} \begin{bmatrix} D_{n_1}u^-(x) \\ \\ D_{t_1}u^-(x) \end{bmatrix} , \tag{2.4.13}$$

Then, using Taylor expansions and the fact that u^- is H^2 on each $e_i \cap \tau^-$, we obtain the desired estimate. □

Before we move on to the definition of the multiscale basis functions, in Corollary 2.4.5 we collect the results of Lemmas 2.4.2, 2.4.3 and 2.4.4 in a simpler form. To do this, we introduce the 6×6 matrix

$$M_{\widehat{\alpha},\theta_1,\theta_2,\beta} := \begin{bmatrix} I & 0 & -A_{\widehat{\alpha},\theta_1} \\ 0 & I & -A_{\widehat{\alpha},\theta_2} R_{\theta_2-\theta_1-\beta} \\ \mathcal{R}_1 & \mathcal{R}_2 & 0 \end{bmatrix} ,$$

where

$$\mathcal{R}_1 = \begin{bmatrix} r_1^- & r_1^+ \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathcal{R}_2 = \begin{bmatrix} 0 & 0 \\ r_2^- & r_2^+ \end{bmatrix} . \tag{2.4.14}$$

Also, for each $v \in H_0^1(\Omega)$ with suitably well-defined point values at $y_i,\ i = 1, 2$, we define the vectors $\mathbf{c}(v)\ \mathbf{d}(v) \in \mathbb{R}^6$ by

$$\mathbf{c}(v) = [0, 0, 0, 0, v(x_1) - v(x_3), v(x_2) - v(x_3)]^T , \tag{2.4.15}$$

and

$$\mathbf{d}(v) := [(D_{e_1} v^-)(y_1), (D_{e_1} v^+)(y_1), (D_{e_2} v^-)(y_2), (D_{e_2} v^+)(y_2), (D_{n_1} v^-)(y_1), (D_{t_1} v^-)(y_1)]^T . \tag{2.4.16}$$

Note that $\mathbf{c}(v)$ and $\mathbf{d}(v)$ depend linearly on v. Then we have

Corollary 2.4.5. *If u is the exact solution of the problem* (2.1.2)*, then for each element τ which intersects the interface as in Assumption 2.4.1, and with the notation defined there, we have*

$$M_{\widehat{\alpha},\theta_1,\theta_2,\beta}\ \mathbf{d}(u) = \mathbf{c}(u) + \boldsymbol{\delta} ,$$

where $\boldsymbol{\delta} \in \mathbb{R}^6$ is defined by

$$\boldsymbol{\delta} \ = \ \begin{bmatrix} \mathbf{0} \\ A_{\widehat{\alpha},\theta_2} \boldsymbol{\varepsilon}' \\ \boldsymbol{\varepsilon} \end{bmatrix} , \tag{2.4.17}$$

and $\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}'$ are as defined in Lemmas 2.4.2 and 2.4.4.

Proof. This is obtained by writing down: (i) Lemma 2.4.3 for $i = 1$; (ii) Lemma 2.4.3 for $i = 2$ combined with Lemma 2.4.4 and (iii) Lemma 2.4.2. □

2.4.2 Novel Interior Boundary Condition and Boundary Error

We now use the relations derived in the previous subsection to derive suitable boundary conditions for multiscale basis functions.

For any element $\tau \in \mathcal{T}_h$, let $x_p,\ p = 1, 2, 3$ denote its nodes. The multiscale basis functions Φ_p^{MS} are found as solutions to the subgrid problems (2.2.1) on τ, subject to Dirichlet boundary data $\phi_{p,\tau}$ on $\partial\tau$ which has to be specified, subject to the nodal condition:

$$\phi_{p,\tau}(x_q) = \delta_{p,q}\ , \quad p, q \in \{1, 2, 3\} \tag{2.4.18}$$

(see (2.2.2)). If the interface Γ does not intersect τ, then we choose $\phi_{p,\tau}$ on $\partial\tau$ to be the linear interpolant of (2.4.18) on each edge of τ, and then the solution of (2.2.1) is also linear on τ. Otherwise (under Assumption 2.4.1), our construction for $\phi_{p,\tau}$ (described below) will be continous on $\partial\tau$, linear on each of the intersected edge segments, $\{\overline{x_3 y_i}, \quad \overline{y_i x_i}, \quad \text{for} \quad i = 1, 2\}$ and linear on the third edge $\overline{x_1 x_2}$. Because of (2.4.18), it remains to specify the gradient of $\phi_{p,\tau}$ on each of the two pieces of the intersected edges. These gradients are computed by Algorithm 2.4.7 below. This requires solving two 6×6 linear systems with the same coefficient matrices. Before we give the algorithm we first establish the solvability of these systems and obtain bounds on their solution which will be needed later.

Theorem 2.4.6. *Under Assumption 2.4.1, suppose* $\phi := \theta_2 - \theta_1 - \beta \neq 0$ *and introduce the* 2×2 *matrix*

$$D \ := \ \mathcal{R}_1 A_{\widehat{\alpha},\theta_1} \ + \ \mathcal{R}_2 A_{\widehat{\alpha},\theta_2} R_\phi.$$

Then, for all $\widehat{\alpha}$ *sufficiently large,* D *is nonsingular, and* $M_{\widehat{\alpha},\theta_1,\theta_2,\beta}$ *is nonsingular with*

$$(M_{\widehat{\alpha},\theta_1,\theta_2,\beta})^{-1} \ = \ \begin{bmatrix} I & 0 & A_{\widehat{\alpha},\theta_1} \\ 0 & I & A_{\widehat{\alpha},\theta_2} R_\phi \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -D^{-1}\mathcal{R}_1 & -D^{-1}\mathcal{R}_2 & D^{-1} \end{bmatrix}. \tag{2.4.19}$$

Moreover

$$\|D^{-1}\|_\infty \ \lesssim \ \widehat{\alpha}^{-1} h_\tau^{-1} (\sin\phi)^{-1}. \tag{2.4.20}$$

Proof. A tedious but elementary calculation shows that with

$$E \;:=\; \begin{bmatrix} r_1^+ \cos\theta_1 & 0 \\ \\ r_2^+ \cos\theta_2 \cos\phi & -r_2^+ \cos\theta_2 \sin\phi \end{bmatrix} ,$$

we have $\|\widehat{\alpha}^{-1}D - E\|_\infty \;\leq\; C\widehat{\alpha}^{-1}h_\tau$, where the constant C is independent of θ_1, θ_2, ϕ, β and h_τ. Since E is non-singular, standard matrix perturbation theory shows that, for large enough $\widehat{\alpha}$,

$$\|\widehat{\alpha}D^{-1}\|_\infty \;=\; \|(\widehat{\alpha}^{-1}D)^{-1}\|_\infty \leq\; C'\|E^{-1}\|_\infty$$

with C' also independent of the above parameters. In fact

$$E^{-1} \;=\; \begin{bmatrix} (r_1^+ \cos\theta_1)^{-1} & 0 \\ \\ -(r_1^+ \cos\theta_1 \sin\phi)^{-1} \cos\phi & -(r_2^+ \cos\theta_2 \sin\phi)^{-1} \end{bmatrix} ,$$

and so (2.4.20) follows directly, recalling Assumption 2.4.1. Because D^{-1} exists, the formula for M^{-1} is verified by simple matrix manipulation. □

This now leads us to Algorithm 2.4.7 for computing the boundary data $\phi_{p,\tau}$ for the multiscale basis functions Φ_p^{MS} on τ.

Algorithm 2.4.7. *For $p = 1, 2, 3$,*

1. Solve the linear system:

$$M_{\widehat{\alpha},\theta_1,\theta_2,\beta}\; \mathbf{d}_p = \mathbf{c}(\phi_{p,\tau}) \;. \tag{2.4.21}$$

2. Then set

$$\begin{cases} (D_{e_1}\phi_{p,\tau})|_{\overline{x_3y_1}} = (\mathbf{d}_p)_1, & (D_{e_1}\phi_{p,\tau})|_{\overline{y_1x_1}} = (\mathbf{d}_p)_2 \;, \\ \\ (D_{e_2}\phi_{p,\tau})|_{\overline{x_3y_2}} = (\mathbf{d}_p)_3, & (D_{e_2}\phi_{p,\tau})|_{\overline{y_2x_2}} = (\mathbf{d}_p)_4 \;. \end{cases} \tag{2.4.22}$$

Remark 2.4.8. (i) The right hand side $\mathbf{c}(\phi_{p,\tau})$ in system (2.4.21) is determined by (2.4.15) and (2.4.18). It is easy to see that $\mathbf{c}(\phi_{1,\tau}) + \mathbf{c}(\phi_{2,\tau}) + \mathbf{c}(\phi_{3,\tau}) = \mathbf{0}$, so

$\mathbf{d}_1 + \mathbf{d}_2 + \mathbf{d}_3 = \mathbf{0}$ and only two of the three systems (2.4.21) has to be solved. Moreover since the function $\phi_{1,\tau} + \phi_{2,\tau} + \phi_{3,\tau}$ has value 1 at each node of τ and zero derivative along each $\overline{x_3 y i}$, $\overline{y_i x_i}$, $i = 1, 2$ and along $\overline{x_1 x_2}$, it thus satisfies the requirements of (2.2.2) and (2.2.3).

(ii) Since $\phi_{p,\tau}$ is defined to be linear on each $e_i \cap \tau^-$ and $e_i \cap \tau^+$ and to satisfy the nodal condition (2.4.18), the continuity of $\phi_{p,\tau}$ at each intersection point y_i is guaranteed by the last two equations in (2.4.21).

(iii) If $\theta_i = 0$ (i.e. the interface intersects edge e_i orthogonally) then the boundary condition computed by Algorithm 2.4.7 coincides with the "oscillatory boundary condition" proposed in [51]. More precisely, if $\theta_1 = 0$, it is easy to see that the first two equations and last two equations of (2.4.21) imply

$$(\mathbf{d}_p)_2 \;=\; \widehat{\alpha}(\mathbf{d}_p)_1 \quad \text{and} \quad \mathrm{r}_1^-(\mathbf{d}_p)_1 + r_1^+(\mathbf{d}_p)_2 \;=\; \phi_{p,\tau}(x_1) - \phi_{p,\tau}(x_3)$$

and hence

$$(\mathbf{d}_p)_1 = \frac{\phi_{p,\tau}(x_1) - \phi_{p,\tau}(x_3)}{r_1^- + \widehat{\alpha}\, r_1^+}, \quad (\mathbf{d}_p)_2 = \widehat{\alpha}\frac{\phi_{p,\tau}(x_1) - \phi_{p,\tau}(x_3)}{r_1^- + \widehat{\alpha}\, r_1^+}. \tag{2.4.23}$$

Thus $\phi_{p,\tau}$ is the solution of the reduced elliptic differential equation $-(\alpha\, \phi'_{p,\tau})' = 0$ on $\overline{x_3 x_1}$, which is exactly how the "oscillatory" boundary condition is constructed.

(iv) When $\theta_i \neq 0$ for $i = 1, 2$ the boundary condition on each e_i depends on *both* θ_1 and θ_2. In particular, if $p = 1$ (respectively 2), the function $\phi_{p,\tau}$ does not necessarily vanish on the edge e_2 (respectively e_1).

(v) Algorithm 2.4.7 determines $\phi_{p,\tau}$ and hence Φ_p^{MS} on each τ individually and does not guarantee that Φ_p^{MS} will be continuous across element edges, so that approximation in $\mathrm{span}\{\Phi_p^{\mathrm{MS}} : p \in \mathscr{I}_h(\Omega)\}$ may not be conforming. We resolve this issue later by averaging across element edges (see Section 2.4.4).

In the next theorem we show that the nodal interpolant $\mathcal{I}_h^{\mathrm{MS}} u = \sum_p u(p)\phi_{p,\tau}$ is a good approximation to u along the boundary of the element τ. Recall the notation $E_h^{\mathrm{MS}} := u - \mathcal{I}_h^{\mathrm{MS}} u$.

Theorem 2.4.9. *Let u be the exact solution of* (2.1.2). *Consider an element τ which intersects the interface as in Assumption 2.4.1. Suppose also $\phi := \theta_2 - \theta_1 - \beta \neq 0$. Then we have, for $m = 0, 1$*

$$\begin{aligned} &\max_{i=1,2} \left\{ \widehat{\alpha} h_\tau^m \|D_{e_i}^m E_h^{\mathrm{MS}}\|_{L_\infty(e_i \cap \tau^-)} \ , \ h_\tau^m \|D_{e_i}^m E_h^{\mathrm{MS}}\|_{L_\infty(e_i \cap \tau^+)} \right\} \\ &\quad \lesssim \ h_\tau^{3/2} \max_{\substack{i=1,2 \\ |k|=1}} \left[\widehat{\alpha}^2 \|D^k D_{e_i} u\|^2_{L_2(e_i \cap \tau^-)} \ + \|D^k D_{e_i} u\|^2_{L_2(e_i \cap \tau^+)} \right]^{1/2} . \quad (2.4.24) \end{aligned}$$

The hidden constant may blow up if $\phi \to 0$. (See Remark 2.4.10 below.)

Proof. We give the proof on the assumption that $\tau^\pm$ are as depicted in Figure 2.4 (i.e. α is large in the region containing x_3 and small in the region containing x_1, x_2). Making use of (2.4.21) and the fact that $\mathbf{c}(u)$ depends only on the nodal values of u, we have

$$M_{\widehat{\alpha},\theta_1,\theta_2,\beta} \left(\sum_{p=1}^{3} u(x_p) \mathbf{d}_p \right) \ = \ \mathbf{c}(\mathcal{I}_h^{\mathrm{MS}} u) \ = \ \mathbf{c}(u) \ .$$

Combining this with Corollary 2.4.5 we obtain

$$M_{\widehat{\alpha},\theta_1,\theta_2,\beta} \left(\mathbf{d}(u) - \sum_{p=1}^{3} u(x_p) \mathbf{d}_p \right) \ = \ \boldsymbol{\delta} \ .$$

Hence, using (2.4.19) and (2.4.17), we obtain

$$\left(\mathbf{d}(u) - \sum_{p=1}^{3} u(x_p) \mathbf{d}_p \right) \ = \ \begin{bmatrix} I & 0 & A_{\widehat{\alpha},\theta_1} \\ 0 & I & A_{\widehat{\alpha},\theta_2} R_\phi \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ A_{\widehat{\alpha},\theta_2} \boldsymbol{\varepsilon}' \\ D^{-1}(\boldsymbol{\varepsilon} - \mathcal{R}_2 A_{\widehat{\alpha},\theta_2} \boldsymbol{\varepsilon}') \end{bmatrix} . \quad (2.4.25)$$

Now by (2.4.16), and (2.4.22), we see that the first four entries of the left-hand side of (2.4.25) are
$D_{e_1}(u - \mathcal{I}_h^{\mathrm{MS}} u)^-(y_1), D_{e_1}(u - \mathcal{I}_h^{\mathrm{MS}} u)^+(y_1), D_{e_2}(u - \mathcal{I}_h^{\mathrm{MS}} u)^-(y_2)$, and $D_{e_2}(u - \mathcal{I}_h^{\mathrm{MS}} u)^+(y_2)$. Examining the right-hand side of (2.4.25), we see that the first two entries are

$$A_{\widehat{\alpha},\theta_1} D^{-1}(\boldsymbol{\varepsilon} - \mathcal{R}_2 A_{\widehat{\alpha},\theta_2} \boldsymbol{\varepsilon}') \ .$$

Now, recalling Lemmas 2.4.2, 2.4.4 and (2.4.20) we obtain

$$\|D^{-1}(\boldsymbol{\varepsilon}-\mathcal{R}_2 A_{\widehat{\alpha},\theta_2}\boldsymbol{\varepsilon}')\|_\infty \lesssim \frac{h_\tau^{1/2}}{\sin\phi}\max_{\substack{i=1,2\\|k|=1}}\left\{\|D^k D_{e_i}u\|^2_{L_2(e_i\cap\tau^-)}+\frac{1}{\widehat{\alpha}^2}\|D^k D_{e_i}u\|^2_{L_2(e_i\cap\tau^+)}\right\}^{1/2}.$$

Hence

$$\begin{aligned}\max\left\{\widehat{\alpha}|D_{e_1}(u-\mathcal{I}_h^{\mathrm{MS}}u)^-(y_1)|,\ |D_{e_1}(u-\mathcal{I}_h^{\mathrm{MS}}u)^+(y_1)|\right\}&\\ \lesssim\ h_\tau^{1/2}\max_{\substack{i=1,2\\|k|=1}}\left[\widehat{\alpha}^2\|D^k D_{e_i}u\|^2_{L_2(e_i\cap\tau^-)}+\|D^k D_{e_i}u\|^2_{L_2(e_i\cap\tau^+)}\right]^{1/2}&.\end{aligned}$$

Similarly, the third and fourth components of (2.4.25) yield the same estimate for $\max\left\{\widehat{\alpha}|D_{e_2}(u-\mathcal{I}_h^{\mathrm{MS}}u)^-(y_2)|,\ |D_{e_2}(u-\mathcal{I}_h^{\mathrm{MS}}u)^+(y_2)|\right\}$.

The estimates (2.4.24) for $m=1$ then readily follow. For example, since $\mathcal{I}_h^{\mathrm{MS}}u$ is linear on $e_i\cap\tau^-$, we have, for $x\in e_i\cap\tau^-$,

$$\begin{aligned}\widehat{\alpha}|D_{e_i}(u-\mathcal{I}_h^{\mathrm{MS}}u)(x)| &\le \widehat{\alpha}|(D_{e_i}u)(x)-(D_{e_i}u^-)(y_i)|+\widehat{\alpha}|D_{e_i}(u-\mathcal{I}_h^{\mathrm{MS}}u)^-(y_i)|\\ &\lesssim h_\tau^{1/2}\max_{\substack{i=1,2\\|k|=1}}\left[\widehat{\alpha}^2\|D^k D_{e_i}u\|^2_{L_2(e_i\cap\tau^-)}+\|D^k D_{e_i}u\|^2_{L_2(e_i\cap\tau^+)}\right]^{1/2}.\end{aligned}\tag{2.4.26}$$

To obtain the estimate for $m=0$, recall that $u-\mathcal{I}_h^{\mathrm{MS}}u$ vanishes at the nodes, so we can write, for $x\in e_1\cap\tau^-$,

$$(u-\mathcal{I}_h^{\mathrm{MS}}u)(x) = \int_{x_3}^x D_{e_1}(u-\mathcal{I}_h^{\mathrm{MS}}u)(t)dt,\tag{2.4.27}$$

and the required estimates for $\|u-\mathcal{I}_h^{\mathrm{MS}}u\|_{L_\infty(e_1\cap\tau^-)}$ follow directly. The remainder of the estimates (2.4.24) for $m=0$ are similar.

□

Remark 2.4.10. The critical case $\phi=\theta_2-\theta_1-\beta=0$ in Theorem 2.4.9 occurs when the unit outward normals $\mathbf{n}_1$ and $\mathbf{n}_2$ to Γ at the two intersection points $\mathbf{y}_1,\ \mathbf{y}_2$ coincide. In this case, if the interface Γ is not a straight line, then τ may be subdivided into two sub-elements, in each of which ϕ no longer vanishes and Algorithm 2.4.7 applies to each of these sub-elements.

However if $\Gamma\cap\tau$ is a straight line, no such refinement will succeed. Instead (referring to the geometry in Fig. 2.4), one may simply subdivide the quadrilateral

τ^+ into two triangles and combine this with τ^- to yield a new mesh which locally resolves Γ and then discretise using standard linear basis functions on each of these three elements.

An alternative approach is suggested by the "Immersed Finite Element (IFE) method" of Li, Lin and Wu [68], where in any case the interface segment $\tau \cap \Gamma$ is approximated by a straight line and a special finite element basis $\Psi_{p,\tau} : p = 1, 2, 3$ is constructed on τ which is required to be affine on each of τ^- and τ^+ and to satisfy the six conditions (with the same geometry as in Assumption 2.4.1):

$$\Psi_{p,\tau}(x_q) = \delta_{p,q} , \quad q = 1, 2, 3 , \tag{2.4.28}$$

$$\Psi^-_{p,\tau}(y_i) = \Psi^+_{p,\tau}(y_i) , \quad i = 1, 2 , \tag{2.4.29}$$

$$\text{and} \quad D_n \Psi^-_{p,\tau} = \widehat{\alpha} D_n \Psi^+_{p,\tau} , \tag{2.4.30}$$

where $\mathbf{n}$ denotes the (constant) normal direction to the straight line $\Gamma \cap \tau$ pointing from τ^- to τ^+ and $\mathbf{t}$ the corresponding tangential direction (as in Fig. 2.4). Note that in (2.4.30), the quantities on each side of the equation are constant since the $\Psi^{\pm}_{p,\tau}$ are assumed affine.

The following lemma shows that the immersed finite element algorithm defines a solution to (2.4.21) even when $\Gamma \cap \tau$ is a straight line (so that $\phi = 0$). However the error estimates of Theorem 2.4.9 are no longer true in general for the IFE approach, as the following example shows.

Example 2.4.11. Consider an element τ with vertices (0, 0), (0, h), (h, h) and the interface is the segment connected by (0, $h/2$), ($h/2$, $h/2$). If we consider $\alpha = \widehat{\alpha}$ on τ^-(left part) and $\alpha = 1$ on τ^+(right part), then the IFE nodal basis function with value 1 at (h, 0) is given by

$$\frac{2\,x}{(\widehat{\alpha}+1)h} - \frac{y}{h}, \quad \text{on} \quad \tau^- \qquad \text{and} \qquad \frac{2\,\widehat{\alpha}\,(x-h)}{(\widehat{\alpha}+1)h} - \frac{y-h}{h}, \quad \text{on} \quad \tau^+ .$$

The directional derivative along the edge from (0, 0), (h, h) in τ^- is $\frac{\sqrt{2}}{(\widehat{\alpha}+1)h} - \frac{1}{\sqrt{2}h}$ of $\mathcal{O}(1/h)$ but not $\mathcal{O}(1/(\widehat{\alpha}\,h))$. Hence the estimate in Theorem 2.4.9 is not true.

Lemma 2.4.12. *Suppose $\Gamma \cap \tau$ is a straight line segment and suppose $\Psi_{p,\tau}$ is defined via* (2.4.28)*,* (2.4.29) *and* (2.4.30)*. Then the vector $\mathbf{d}(\Psi_{p,\tau})$ (defined as in* (2.4.16)*) provides a solution to system* (2.4.21)*.*

Proof. Using the fact that $\Psi_{p,\tau}$ is affine on each of $\tau^{\pm}$, and using also (2.4.29), we have, for $i = 1, 2$,

$$\begin{aligned}(\nabla\Psi^-_{p,\tau})^T(y_i - x_3) \;+\; (\nabla\Psi^+_{p,\tau})^T(x_i - y_i) \quad &= \quad \Psi^-_{p,\tau}(y_i) - \Psi_{p,\tau}(x_3) + \Psi_{p,\tau}(x_i) - \Psi^+_{p,\tau}(y_i) \\ &= \quad \Psi_{p,\tau}(x_i) - \Psi_{p,\tau}(x_3)\ .\end{aligned}$$

Since $r_i^-\mathbf{e}_i = (y_i - x_3)$ and $r_i^+\mathbf{e}_i = (x_i - y_i)$, and since the gradients $\nabla\Psi^{\pm}_{p,\tau}$ are constant, it follows that

$$r_i^- D_{e_i}\Psi^-_{p,\tau}(y_i) + r_i^+ D_{e_i}\Psi^+_{p,\tau}(y_1) \;=\; \Psi_{p,\tau}(x_i) - \Psi_{p,\tau}(x_3) \quad \text{for} \quad i = 1, 2\ ,$$

and so the last two equations of (2.4.21) are satisfied.

By a similar argument,

$$(\nabla\Psi^-_{p,\tau})^T(y_1 - y_2) \;=\; \Psi^-_{p,\tau}(y_1) - \Psi^-_{p,\tau}(y_2) \;=\; \Psi^+_{p,\tau}(y_1) - \Psi^+_{p,\tau}(y_2) \;=\; (\nabla\Psi^+_{p,\tau})^T(y_1 - y_2)\ ,$$

and since $y_1 - y_2$ is in the direction of $\mathbf{t}$ (the tangent direction along Γ), this implies

$$D_t\Psi^+_{p,\tau} = D_t\Psi^-_{p,\tau} \quad \text{on} \quad \Gamma \cap \tau\ .$$

Combining this with (2.4.2) and (2.4.30), we have

$$\begin{aligned}D_{e_1}\Psi^-_{p,\tau} \quad &= \quad \cos\theta_1 D_n\Psi^-_{p,\tau} \;+\; \sin\theta_1 D_t\Psi^-_{p,\tau} \\ D_{e_1}\Psi^+_{p,\tau} \quad &= \quad \widehat{\alpha}\cos\theta_1 D_n\Psi^-_{p,\tau} \;+\; \sin\theta_1 D_t\Psi^-_{p,\tau}\ ,\end{aligned}$$

and thus the first two equations in (2.4.21) are satisfied. The verification of the third and fourth equations in (2.4.21) is entirely analogous.

□

Remark 2.4.13. The previous lemma shows that the system (2.4.21) is consistent when $\Gamma \cap \tau$ is a straight line, which is a particular case of $\phi = 0$. Under the general assumption only that $\phi = 0$, and examining the proof of Theorem 2.4.6 we see that in this case $D = \mathcal{R}_1 A_{\widehat{\alpha},\theta_1} + \mathcal{R}_2 A_{\widehat{\alpha},\theta_2}$. The (non)singularity of this for general choices of $\widehat{\alpha}$, and θ_i, r_i^-, r_i^+, $i = 1, 2$ has not yet been analysed.

2.4.3 Interior Error

The main result in this section is Theorem 2.4.16, which gives an $\alpha-$ explicit estimate for the error $|u-u_h^{\mathrm{MS}}|_{H^1(\tau),\alpha}$, in the case where the interface may cut through τ. This is obtained by an application of Theorem 2.2.2, and thus requires we first show that $E_h^{\mathrm{MS}} = u - \mathcal{I}_h^{\mathrm{MS}} u$ can be extended from the boundary to the interior of τ in a suitably robust way. This extension is proved in Lemma 2.4.15, which requires a further technical assumption on the geometry of $\Gamma \cap \tau$.

Assumption 2.4.14. We impose Assumption 2.4.1 and further assume that when Γ intersects any element τ, $\Gamma \cap \tau$ is star shaped about x_3. That is, introducing polar coordinates with origin x_3 and polar angle θ measured anticlockwise from e_1, we assume that each $(x,y) \in \Gamma \cap \tau$ can be written $(x,y) = (r(\theta)\cos\theta, r(\theta)\sin\theta)$, for $\theta \in [0,\beta]$. Writing also the edge $\overline{x_1 x_2}$ as $(x,y) = (r^*(\theta)\cos\theta, r^*(\theta)\sin\theta)$ for a suitable function r^*, we assume there exist constants $C > 0$ and, $1 > C^* > 0$ such that

$$|r'(\theta)| \leq C r(\theta) \quad \text{and} \quad r(\theta) \leq C^* r^*(\theta)\ , \quad \text{for all} \quad \theta \in [0,\beta]. \tag{2.4.31}$$

Note that under this assumption, we can integrate the left-hand side of (2.4.31) to obtain $|\log(r(\theta)/r(0))| \leq C\beta$, and since $r(0) = r_1^-$, we can combine this with (2.4.1) to obtain

$$r(\theta) \sim h_\tau \qquad \text{for all} \quad \theta \in [0,\beta]\ . \tag{2.4.32}$$

Now letting s denote arclength along $\Gamma \cap \tau$, it is easily seen that

$$ds = \sqrt{(r(\theta))^2 + (r'(\theta))^2}\, d\theta \sim h_\tau d\theta. \tag{2.4.33}$$

Moreover, since (2.4.31) implies $r^*(\theta) \sim h_\tau$, we also have

$$|\tau^+| = \int_0^\beta \int_{r(\theta)}^{r_1(\theta)} r\, dr d\theta = \frac{1}{2}\int_0^\beta \left[(r^*)^2(\theta) - r^2(\theta)\right]\, d\theta \sim h_\tau^2\ .$$

A similar but simpler argument shows $|\tau^-| \sim h_\tau^2$. Collecting these relations, we have

$$|\Gamma \cap \tau| \sim h_\tau\ , \qquad |\tau^\pm| \sim h_\tau^2\ . \tag{2.4.34}$$

These are needed in the proof of the following result.

Lemma 2.4.15. *Under Assumption 2.4.14 there exists* $\widetilde{E}_h^{\mathrm{MS}} \in H^1(\tau)$ *with* $\widetilde{E}_h^{\mathrm{MS}} = E_h^{\mathrm{MS}}$ *on* $\partial\tau$ *and satisfying*

$$|\widetilde{E}_h^{\mathrm{MS}}|^2_{H^1(\tau),\alpha} \;\lesssim\; h_\tau^2 \left(\widehat{\alpha} \max_{i=1,2,3} \|D_{e_i} E_h^{\mathrm{MS}}\|^2_{L_\infty(e_i \cap \tau^-)} \;+\; \max_{i=1,2,3} \|D_{e_i} E_h^{\mathrm{MS}}\|^2_{L_\infty(e_i \cap \tau^+)} \right). \tag{2.4.35}$$

Proof. For notational convenience in the proof we make the abbreviations: $E = E_h^{\mathrm{MS}}$, $\widetilde{E} = \widetilde{E}_h^{\mathrm{MS}}$. We assume the geometric situation as in Figure 2.4, so that τ^- (the region where α is high) contains the node x_3. The case where τ^- contains two nodes is entirely analogous.

Using Assumption 2.4.14, we parameterize τ^- by introducing local coordinates (t, θ) such that

$$x = t\, r(\theta) \cos\theta, \qquad y = t\, r(\theta) \sin\theta, \qquad t \in [0,1]\ , \quad \theta \in [0, \beta]\ . \tag{2.4.36}$$

Then we define $\widetilde{E}$ on τ^- explicitly by:

$$\widetilde{E}(t,\theta) \;=\; \left(\frac{\theta}{\beta}\right) E(x_3 + t r_2^- \mathbf{e}_2) \;+\; \left(1 - \frac{\theta}{\beta}\right) E(x_3 + t r_1^- \mathbf{e}_1)\ , \quad t \in [0,1]\ , \quad \theta \in [0,\beta]\ . \tag{2.4.37}$$

Clearly $\widetilde{E}$ coincides with E on $e_i \cap \tau^-$, for each $i = 1, 2$ and, moreover,

$$\begin{aligned} \frac{\partial \widetilde{E}}{\partial x}(t,\theta) \;\;=\;\; & \left(\left(\frac{\theta}{\beta}\right) r_2^-(D_{e_2}E)(x_3 + t r_2^- \mathbf{e}_2) \;+\; \left(1 - \frac{\theta}{\beta}\right) r_1^-(D_{e_1}E)(x_3 + t r_1^- \mathbf{e}_1) \right) \frac{\partial t}{\partial x} \\ & + \frac{1}{\beta} \Big(E(x_3 + t r_2^- \mathbf{e}_2) - E(x_3 + t r_1^- \mathbf{e}_1) \Big) \frac{\partial \theta}{\partial x} \end{aligned} \tag{2.4.38}$$

with an analogous formula for $\partial\widetilde{E}/\partial y$. Defining the Jacobian

$$J := \begin{bmatrix} \frac{\partial x}{\partial t} & \frac{\partial x}{\partial \theta} \\ \\ \frac{\partial y}{\partial t} & \frac{\partial y}{\partial \theta} \end{bmatrix}\ , \quad \text{we have} \quad \det(J) \;=\; t\, r^2(\theta) \tag{2.4.39}$$

and (abbreviating $r(\theta)$ by r), the partial derivatives of θ and t are computed:

$$\begin{bmatrix} \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \\ \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{bmatrix} = J^{-1} \;=\; \frac{1}{t\, r^2} \begin{bmatrix} t(r\cos\theta + r'\sin\theta) & t(r\sin\theta - r'\cos\theta) \\ -r\sin\theta & r\cos\theta \end{bmatrix}. \tag{2.4.40}$$

Making use of (2.4.40) and then Assumption 2.4.14 (which includes Assumption 2.4.1), the first term on the right-hand side of (2.4.38) is estimated by

$$\frac{h_\tau}{r(\theta)}\ \left|\cos\theta+\frac{r'(\theta)}{r(\theta)}\sin\theta\right|\ \max_{i=1,2}\|D_{e_i}E\|_{L_\infty(e_i\cap\tau^-)}\ \lesssim\ \frac{h_\tau}{r(\theta)}\ \max_{i=1,2}\|D_{e_i}E\|_{L_\infty(e_i\cap\tau^-)}\ . \tag{2.4.41}$$

Moreover since $E(x_3)=0$, we have $|E(x_3+tr_i^-e_i)|\ \lesssim\ th_\tau\|D_{e_i}E\|_{L_\infty(e_i\cap\tau^-)}$, $i=1,2$, and hence the second term on the right-hand side of (2.4.38) is bounded exactly as in the right-hand side of (2.4.41). An analogous procedure can be applied to $\partial\widetilde{E}/\partial y$, thus yielding, overall,

$$|\nabla\widetilde{E}(t,\theta)|\ \lesssim\ \frac{h_\tau}{r(\theta)}\ \max_{i=1,2}\|D_{e_i}E\|_{L_\infty(e_i\cap\tau^-)}\quad\text{for}\quad t\in[0,1],\ \theta\in[0,\beta]\ .$$

Therefore, using also (2.4.39), we obtain the estimate on τ^-:

$$\begin{aligned}|\widetilde{E}|^2_{H^1(\tau^-),\alpha}\ =\ \int_{\tau_-}\widehat{\alpha}|\nabla\widetilde{E}(x,y)|^2\,dx\,dy\quad&=\quad\widehat{\alpha}\int_0^1\int_0^\beta|\nabla\widetilde{E}(t,\theta)|^2\,t\,r^2(\theta)\,d\theta\,dt\\&\lesssim\quad h_\tau^2\ \widehat{\alpha}\ \max_{i=1,2}\|D_{e_i}E\|^2_{L_\infty(e_i\cap\tau^-)}\ .\end{aligned}\tag{2.4.42}$$

Note that we constructed above an explicit expansion $\widetilde{E}$ into τ^- whose precise behaviour is quite delicate. For the extension into τ^+, it turns out to be sufficient to apply the inverse trace theorem, which only obtains the extension implicitly. Since τ^+ is a Lipschitz domain, the (inverse) trace theorem (using also (2.4.34)), gives an extension $\widetilde{E}$ which satisfies (since $\alpha\lesssim 1$ on τ^+),

$$\begin{aligned}|\widetilde{E}|^2_{H^1(\tau^+),\alpha}\quad&\lesssim\quad|\widetilde{E}|^2_{H^1(\tau^+)}\\&\lesssim\quad h_\tau{}^{-1}\|\widetilde{E}\|^2_{L_2(\partial\tau^+)}\ +\ h_\tau|\widetilde{E}|^2_{H^1(\partial\tau^+)}\\&=\quad\sum_{i=1,2,3}\Big(h_\tau{}^{-1}\|E\|^2_{L_2(e_i\cap\tau^+)}\ +\ h_\tau|E|^2_{H^1(e_i\cap\tau^+)}\Big)\\&\qquad\qquad+\ h_\tau{}^{-1}\|\widetilde{E}\|^2_{L_2(\Gamma\cap\tau)}\ +\ h_\tau|\widetilde{E}|^2_{H^1(\Gamma\cap\tau)}\\&\lesssim\quad h_\tau^2\ \max_{i=1,2,3}\|D_{e_i}E\|^2_{L_\infty(e_i\cap\tau^+)}\ +\ h_\tau{}^{-1}\|\widetilde{E}\|^2_{L_2(\Gamma\cap\tau)}\ +\ h_\tau|\widetilde{E}|^2_{H^1(\Gamma\cap\tau)},\end{aligned}\tag{2.4.43}$$

where in the final estimate we used $E(x_1)=0=E(x_2)$.

It remains to estimate the final two terms in (2.4.43). First note that on $\Gamma\cap\tau$, we

can write $\widetilde{E} = \widetilde{E}(1,\theta)$, for some $\theta \in [0,\beta]$ and by (2.4.37) we have (as above),

$$\|\widetilde{E}\|_{L_\infty(\Gamma\cap\tau)} \lesssim \max_{i=1,2}|E(y_i)| \lesssim h_\tau \max_{i=1,2}\|D_{e_i}E\|_{L_\infty(e_i\cap\tau^+)} . \tag{2.4.44}$$

Moreover, writing $\theta = \theta(s)$ where s denotes arclength along $\Gamma\cap\tau$, we have

$$\left|\frac{d}{ds}\{\widetilde{E}(1,\theta(s))\}\right| = \frac{1}{\beta}|E(y_2)-E(y_1)|\left|\frac{d\theta}{ds}\right| \lesssim h_\tau \max_{i=1,2}\|D_{e_i}E\|_{L_\infty(e_i\cap\tau^+)}\left|\frac{d\theta}{ds}\right| .$$

Hence, making use of (2.4.33) and (2.4.34), we have

$$\begin{aligned} |\widetilde{E}|^2_{H^1(\Gamma\cap\tau^+)} &\lesssim h_\tau^2 \max_{i=1,2}\|D_{e_i}E\|^2_{L_\infty(e_i\cap\tau^+)} \int_0^{|\Gamma\cap\tau|}\left|\frac{d\theta}{ds}\right|^2 ds \\ &\sim h_\tau \max_{i=1,2}\|D_{e_i}E\|^2_{L_\infty(e_i\cap\tau^+)}. \end{aligned} \tag{2.4.45}$$

The Lemma follows on insertion of (2.4.44) and (2.4.45) into (2.4.43).

□

Theorem 2.4.16. *Let u be the solution of (2.1.2) and suppose τ is one of the elements which are cut through by the interface Γ. Then, under Assumption 2.4.14,*

$$\begin{aligned} |E_h^{\mathrm{MS}}|^2_{H^1(\tau),\alpha} &\lesssim h_\tau^2\,\widehat{\alpha}^2\left[|u|^2_{H^2(\tau^-)} + h_\tau|u|^2_{H^{5/2}(\tau^-)}\right] \\ &+ h_\tau^2\left[|u|^2_{H^2(\tau^+)} + h_\tau|u|^2_{H^{5/2}(\tau^+)}\right] + h_\tau^2\,\|f\|^2_{L_2(\tau)}. \end{aligned} \tag{2.4.46}$$

Remark 2.4.17. When τ is an element which is not intersected by the interface, the estimate (2.4.46) still holds but the terms in $|u|_{H^{5/2}(\tau^\pm)}$ are absent.

Proof. From Theorem 2.2.2 and Lemma 2.4.15 we have

$$\begin{aligned} |E_h^{\mathrm{MS}}|^2_{H^1(\tau),\alpha} \lesssim h_\tau^2 \Bigg(&\widehat{\alpha}\max_{i=1,2,3}\|D_{e_i}E_h^{\mathrm{MS}}\|^2_{L_\infty(e_i\cap\tau^-)} \\ &+ \max_{i=1,2,3}\|D_{e_i}E_h^{\mathrm{MS}}\|^2_{L_\infty(e_i\cap\tau^+)} + \|f\|^2_{L_2(\tau)}\Bigg). \end{aligned} \tag{2.4.47}$$

The edge derivatives on the right-hand side of (2.4.47) may be estimated by Theorem 2.4.9, yielding

$$\begin{aligned} |E_h^{\mathrm{MS}}|^2_{H^1(\tau),\alpha} \lesssim\ & h_\tau^3 \max_{\substack{i=1,2,3\\|k|=1}}\left[\widehat{\alpha}^2\|D^kD_{e_i}u\|^2_{L_2(e_i\cap\tau^-)} + \|D^kD_{e_i}u\|^2_{L_2(e_i\cap\tau^+)}\right] \\ &+ h_\tau^2\|f\|^2_{L_2(\tau)} . \end{aligned} \tag{2.4.48}$$

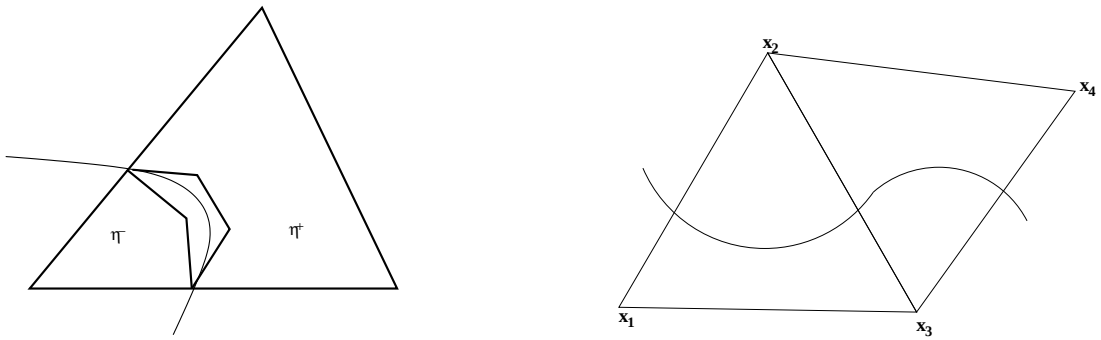


Figure 2.6: Left: Γ is the curved interface and η^- (η^+) is the polygon inside τ^- (respectively τ^+). Right: Two elements intersected by the interface.

Now we adapt the procedure from the proof of Theorem 2.3.1 to bound the terms in u appearing on the right-hand side of (2.4.48). Let η^- be a polygon chosen inside τ^- with the property that $\partial\tau \cap \tau^- \subset \partial\eta^-$ and let η^+ be chosen analogously (see Figure 2.6, left). Clearly we may choose these polygons so that $|\eta^\pm| \sim |\tau^\pm|$.

Then, for $|k| = 1$ and $i = 1, 2$, we have (cf. (2.3.6))

$$\begin{aligned} \|D^k D_{e_i} u^-\|^2_{L_2(e_i\cap\tau^-)} &= |D^k u^-|^2_{H^1(e_i\cap\tau^-)} \\ &\lesssim h_\tau^{-1}|D^k u^-|^2_{H^1(\eta^-)} + |D^k u^-|^2_{H^{3/2}(\eta^-)} \\ &\lesssim h_\tau^{-1}|u^-|^2_{H^2(\tau^-)} + |u^-|^2_{H^{5/2}(\tau^-)}. \end{aligned} \tag{2.4.49}$$

Analogously we have

$$\|D^k D_{e_i} u^+\|^2_{L_2(e_i\cap\tau^+)} \lesssim h_\tau^{-1}|u^+|^2_{H^2(\tau^+)} + |u^+|^2_{H^{5/2}(\tau^+)}. \tag{2.4.50}$$

The required result follows by combining (2.4.49) and (2.4.50) with (2.4.48) □

2.4.4 Conforming Modification and Global Error Estimate

The multiscale basis functions discussed in the previous sections were obtained by solving (2.2.1) on each element τ individually, using a boundary condition relevant to that particular element. When an interface cuts an element edge there is no guarantee that the boundary condition will match across that edge, and so the basis constructed in this way may be discontinuous (i.e. the element may be non-conforming). However, as we now show, it is easy to make the basis functions continuous by local averaging. Consider the interface crossing an edge belonging to two adjacent elements, as in

Fig. 2.6 (right). Let x_p denote any one of the nodes of this pair of triangles and denote the boundary condition on $\overline{x_2x_3}$ (constructed by the method in Section 2.4.2) for $\tau = \triangle x_1x_2x_3$ by $\phi_{p,\tau}$ and the analogous boundary condition for $\tau' = \triangle x_4x_2x_3$ by $\phi_{p,\tau'}$. Then we simply define the averaged boundary condition on $\overline{x_2x_3}$ to be

$$\frac{(\phi_{p,\tau} + \phi_{p,\tau'})}{2} .$$

Doing this for all edges cut by the interface yields a conforming method. Moreover a simple application of the triangle inequality shows that the new boundary condition yields multiscale basis functions and an interpolation operator which satisfies the estimate in Theorem 2.4.9 and hence Theorem 2.4.16 remains true. However, the price we pay is that the resulting basis functions may have a slightly bigger support than the standard linear functions. For example, in Fig. 2.6 (Right), when $p = 4$ the basis function Φ_p^{MS} will not necessarily vanish in the triangle $x_1x_2x_3$.

Theorem 2.4.18. *Suppose Assumption 2.4.14 holds for each element which is cut through by the interface. Suppose also that $f \in H^{1/2}(\Omega)$. Let u be the solution of (2.1.2). Assume also that h is sufficiently small. Then*

$$(i) \quad |u - u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \lesssim h \left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L_2(\Omega)} \right]^{1/2}, \tag{2.4.51}$$

$$(ii) \quad \|u - u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \lesssim h^2 \left[h|f|^2_{H^{1/2}(\Omega)} + \|f\|^2_{L_2(\Omega)} \right]^{1/2}. \tag{2.4.52}$$

Proof. Consider first Case I (see (2.1.5)). By the optimality of u^{MS} in the energy norm, we get

$$|u - u_h^{\mathrm{MS}}|^2_{H^1(\Omega),\alpha} \leq |E_h^{\mathrm{MS}}|^2_{H^1(\Omega),\alpha} = \sum_\tau |E_h^{\mathrm{MS}}|^2_{H^1(\tau),\alpha}, \tag{2.4.53}$$

We now employ Theorem 2.4.16 to estimate the right-hand side. Bearing in mind the regularity estimates in Theorem 2.2.5, since $\partial\Omega$ is assumed to be a convex polygon, we need to assume here that h is small enough that all elements which cross Γ are separated from $\partial\Omega$ by at least a fixed distance. For example, $h \approx \mathrm{dist}(\Gamma, \partial\Omega)/2$ is

sufficient. Under this assumption, and bearing in mind Remark 2.4.17, we have

$$|u-u_h^{\mathrm{MS}}|^2_{H^1(\Omega),\alpha} \lesssim$$
$$h^2\left\{\widehat{\alpha}^2\ \sum_{i=1}^{m}\left(|u|^2_{H^2(\Omega_i)}+h|u|^2_{H^{5/2}(\Omega_i)}\right)\ +\ |u|^2_{H^2(\Omega_0)}+h|u|^2_{H^{5/2}(\Omega_0)}\ +\ \|f\|^2_{L_2(\Omega)}\right\}. \tag{2.4.54}$$

The required estimate (2.4.51) follows from regularity theorem 2.2.5. The estimate (2.4.52) is derived by adapting the duality argument used in the proof of Theorem 2.3.1. The proof in Case II similar. □

Remark 2.4.19. If u is required to satisfy an inhomogeneous boundary condition g (as described in Remark A.2.1) then we have the following estimates for Case I,

$$(i)\quad |u-u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \lesssim h\left[\|f\|^2_{L_2(\Omega)}+\|g\|^2_{H^2(\Omega_0)}\ +h(|f|^2_{H^{1/2}(\Omega)}+|g|^2_{H^{5/2}(\Omega_0)})\right]^{1/2},$$
$$(ii)\quad \|u-u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \lesssim h^2\left[\|f\|^2_{L_2(\Omega)}+\|g\|^2_{H^2(\Omega_0)}\ +h(|f|^2_{H^{1/2}(\Omega)}+|g|^2_{H^{5/2}(\Omega_0)})\right]^{1/2},$$

and, for Case II,

$$(i)\quad |u-u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha} \lesssim h\left[\|f\|^2_{L_2(\Omega)}+\widehat{\alpha}^2\|g\|^2_{H^2(\Omega_0)}\ +h(|f|^2_{H^{1/2}(\Omega)}+\widehat{\alpha}^2|g|^2_{H^{5/2}(\Omega_0)})\right]^{1/2},$$
$$(ii)\quad \|u-u_h^{\mathrm{MS}}\|_{L_2(\Omega)} \lesssim h^2\left[\|f\|^2_{L_2(\Omega)}+\widehat{\alpha}^2\|g\|^2_{H^2(\Omega_0)}\ +h(|f|^2_{H^{1/2}(\Omega)}+\widehat{\alpha}^2|g|^2_{H^{5/2}(\Omega_0)})\right]^{1/2},$$

The latter estimates are pessimistic in some inhomogeneous Dirichlet cases. For example if $u=u_0+C$ where u_0 enjoys the same estimates as in Theorem 2.2.5, then since $\mathcal{I}_h^{\mathrm{MS}}$ preserves constants (see (2.2.5)),

$$|u-u_h^{\mathrm{MS}}|_{H^1(\Omega),\alpha}\ \le\ |u-\mathcal{I}_h^{\mathrm{MS}}u|_{H^1(\Omega),\alpha}\ =\ |u_0-\mathcal{I}_h^{\mathrm{MS}}u_0|_{H^1(\Omega),\alpha} \tag{2.4.55}$$

and the results of Theorem 2.4.18 remain valid in this case.

2.5 Numerical Experiments

In this section, we perform three numerical experiments to verify the convergence rates established above. We consider the weak form of the Dirichlet boundary value

problem:

$$-\nabla \cdot (\alpha \nabla u) = f, \quad \text{in} \quad \Omega, \qquad \text{with} \quad u = g, \quad \text{on} \quad \partial\Omega\,, \tag{2.5.1}$$

for different domains Ω, piecewise constant function α, source term f and boundary condition g. To compute the multiscale basis functions, we subdivide each coarse grid element into M uniform triangular sub-elements and apply the IFE method of [68] for basis function calculation on each coarse element. As discussed in Remark 2.4.10, this involves approximating the interface by a straight line in each element of the fine grid which it intersects. Clearly extending the theoretical results in this paper to this case will require M to be sufficiently large relative to h and α. In the following three experiments, we use $M = 1024$ to ensure that the errors in computing the basis functions are small. At the end of this section we study in more detail how the choice of M affects the overall error in the method.

Experiment 1. In this experiment, $\Omega = [-1,1]\times[-1,1]$, $f = -9\,r$, $g = \frac{r^3}{\alpha_0} + \left(\frac{1}{\alpha_1} - \frac{1}{\alpha_0}\right) r_0^3$, and

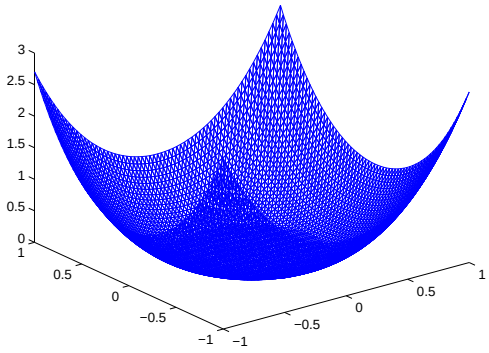
$$\alpha = \begin{cases} \alpha_1 & , \quad r < r_0, \\ \alpha_0 & , \quad r \geq r_0. \end{cases}, \tag{2.5.2}$$

where $r = (x^2+y^2)^{1/2}$ and $r_0 = \pi/6.28$ (see also [68]). The exact solution is

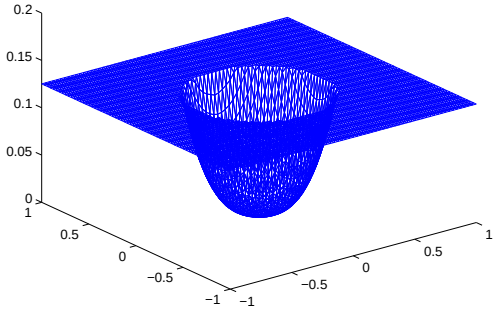
$$u(r,\theta) = \begin{cases} \frac{r^3}{\alpha_1} & , \quad r < r_0, \\ \frac{r^3}{\alpha_0} + \left(\frac{1}{\alpha_1} - \frac{1}{\alpha_0}\right) r_0^3 & , \quad r \geq r_0. \end{cases} \tag{2.5.3}$$

Recalling (2.1.5) and (2.1.6) we shall study Case I: $\alpha_1 = \widehat{\alpha},\ \alpha_0 = 1$ and Case II: $\alpha_1 = 1,\ \alpha_0 = \widehat{\alpha}$. Notice that the source term f is independent of $\widehat{\alpha}$. By examining the form of u and recalling (2.4.55) it follows that our method will enjoy the error estimates of Theorem 2.4.18 in both Cases I and II.

The coarse grid in this case is a uniform triangular grid on Ω. We depict the numerical solutions for both cases in Fig. 2.7. The solutions are flat in the region where the coefficient α is high. Fig. 2.8 shows that the errors are small but are concentrated along the interface. The errors presented in Tables 2.1 and 2.2 show that the method is first order in the H^1 semi-norm and second order in the L_2 norm,

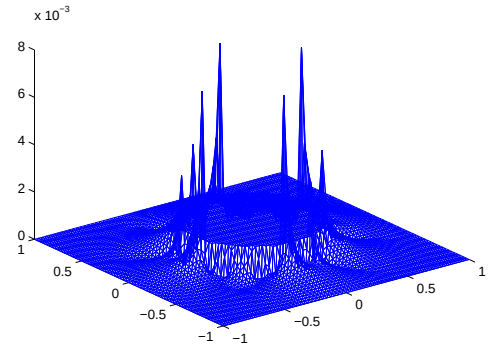


(a) u_h^{MS} for Case I with $\widehat{\alpha} = 100000$.

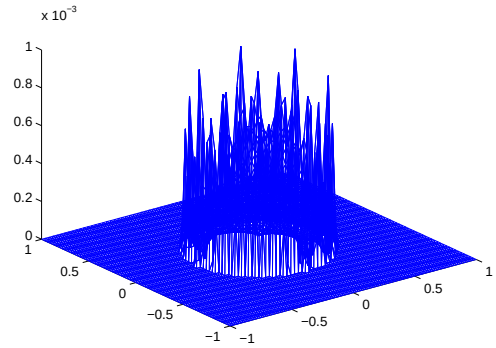


(b) u_h^{MS} for Case II with $\widehat{\alpha} = 100000$.

Figure 2.7: Numerical solutions u_h^{MS} with $h = 1/32$ for Experiment 1.



(a) $|E_h^{\mathrm{MS}}|$ for Case I with $\widehat{\alpha} = 100000$.



(b) $|E_h^{\mathrm{MS}}|$ for Case II with $\widehat{\alpha} = 100000$.

Figure 2.8: Pointwise errors E_h^{MS} and $h = 1/32$ for Experiment 1.

as predicted by the theory.[4] The independence of $\widehat{\alpha}$ can be observed from Figure 2.9.

In this experiment, we can see that our new multiscale finite element method gives much better performance than the standard linear finite element method on the same grid (see Fig 2.10). The improvement is more significant when $\widehat{\alpha}$ is very large, which may be expected since we have proved that our multiscale method converges independently of $\widehat{\alpha}$ whereas the asymptotic constant in the error estimate for the standard finite element method may depend on $\widehat{\alpha}$. For example, when $\widehat{\alpha} = 10^5$, the multiscale finite element method has an L_2 norm error about 66 times smaller than that of the standard linear finite element method, while in the H^1 semi-norm the error is better by a factor of about 15.

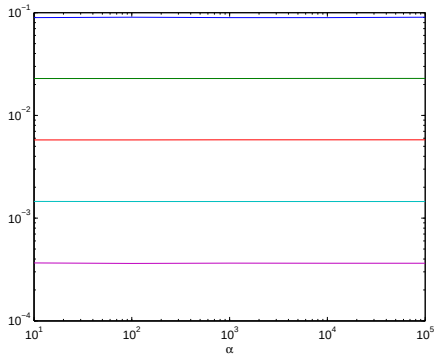
We have also compared our multiscale finite element method with the IFE method

[4]Throughout, we use least squares fitting to estimate the convergence rates.

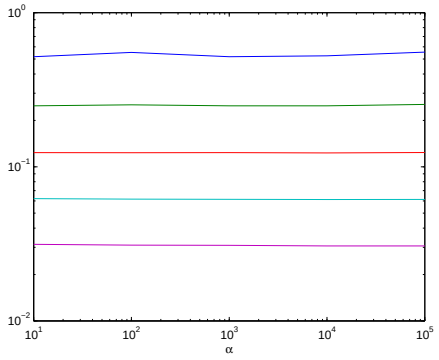
h	$\hat{\alpha} = 10$	$\hat{\alpha} = 100$	$\hat{\alpha} = 1000$	$\hat{\alpha} = 10000$	$\hat{\alpha} = 100000$
1/4	8.9457e-2	9.0295e-2	8.9569e-2	8.9489e-2	9.0375e-2
1/8	2.2833e-2	2.2877e-2	2.2881e-2	2.2891e-2	2.2912e-2
1/16	5.7666e-3	5.7703e-3	5.7791e-3	5.7824e-3	5.7808e-3
1/32	1.4548e-3	1.4521e-3	1.4511e-3	1.4517e-3	1.4511e-3
1/64	3.6619e-4	3.6242e-4	3.6482e-4	3.6369e-4	3.6366e-4
rate	1.9837	1.9899	1.9858	1.9865	1.9895

h	$\widehat{\alpha} = 10$	$\widehat{\alpha} = 100$	$\widehat{\alpha} = 1000$	$\widehat{\alpha} = 10000$	$\widehat{\alpha} = 100000$
1/4	5.1756e-1	5.5251e-1	5.1793e-1	5.2480e-1	5.5458e-1
1/8	2.4868e-1	2.5246e-1	2.4854e-1	2.4858e-1	2.5381e-1
1/16	1.2349e-1	1.2339e-1	1.2355e-1	1.2297e-1	1.2377e-1
1/32	6.2156e-2	6.1687e-2	6.1456e-2	6.1289e-2	6.1355e-2
1/64	3.1374e-2	3.1011e-2	3.0915e-2	3.0651e-2	3.0662e-2
rate	1.0088	1.0343	1.0149	1.0216	1.0402

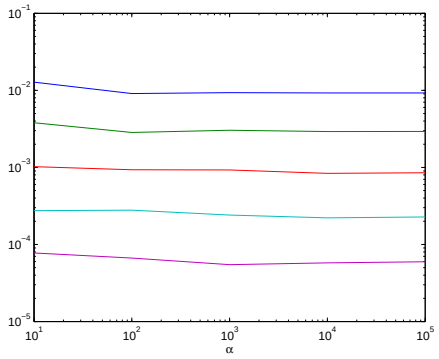
Table 2.1: The L_2-norm errors (upper) and the H^1 semi-norm errors (lower) for the Case I: $\alpha_1 = \widehat{\alpha}$, $\alpha_0 = 1$ in Experiment 1.



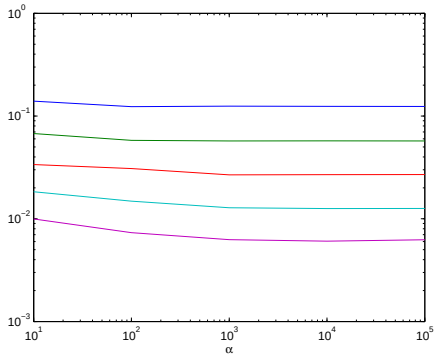
(a) L_2 norm error for Case I.



(b) H_1 semi-norm error for Case I.



(c) L_2 norm error for Case II.



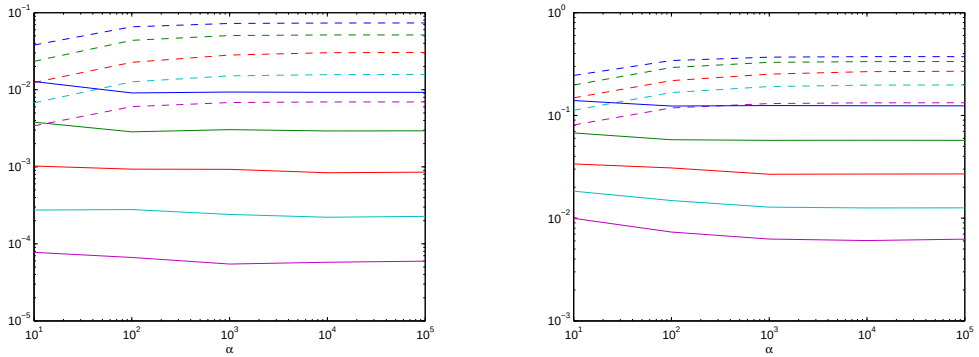
(d) H_1 semi-norm error for Case II.

Figure 2.9: Figures (a)-(d) show that the errors are not affected by the values of $\widehat{\alpha}$ in Experiment 1. Each line represents the error versus $\widehat{\alpha}$ for fixed h. The values of h are 1/4, 1/8, 1/16, 1/32, 1/64 from top to bottom.

h	$\widehat{\alpha}=10$	$\widehat{\alpha}=100$	$\widehat{\alpha}=1000$	$\widehat{\alpha}=10000$	$\widehat{\alpha}=100000$
1/4	1.2782e-2	9.0781e-3	9.3489e-3	9.2490e-3	9.2439e-3
1/8	3.7991e-3	2.8410e-3	3.0394e-3	2.9212e-3	2.9314e-3
1/16	1.0235e-3	9.3213e-4	9.2752e-4	8.3648e-4	8.5214e-4
1/32	2.7485e-4	2.7843e-4	2.4049e-4	2.2169e-4	2.2716e-4
1/64	7.7605e-5	6.6592e-5	5.4716e-5	5.7664e-5	5.9580e-5
rate	1.8517	1.7533	1.8493	1.8371	1.8245

h	$\widehat{\alpha}=10$	$\widehat{\alpha}=100$	$\widehat{\alpha}=1000$	$\widehat{\alpha}=10000$	$\widehat{\alpha}=100000$
1/4	1.3950e-1	1.2346e-1	1.2486e-1	1.2422e-1	1.2408e-1
1/8	6.7497e-2	5.7930e-2	5.7251e-2	5.7320e-2	5.7267e-2
1/16	3.3704e-2	3.0806e-2	2.6738e-2	2.6893e-2	2.6961e-2
1/32	1.8304e-2	1.4854e-2	1.2806e-2	1.2563e-2	1.2609e-2
1/64	9.9543e-3	7.3327e-3	6.2600e-3	6.0577e-3	6.2529e-3
rate	0.9987	0.9708	0.9982	1.0063	1.0160

Table 2.2: The L_2-norm errors (upper) and the H^1 semi-norm errors (lower) for the Case II: $\alpha_1 = 1$, $\alpha_0 = \widehat{\alpha}$ in Experiment 1.



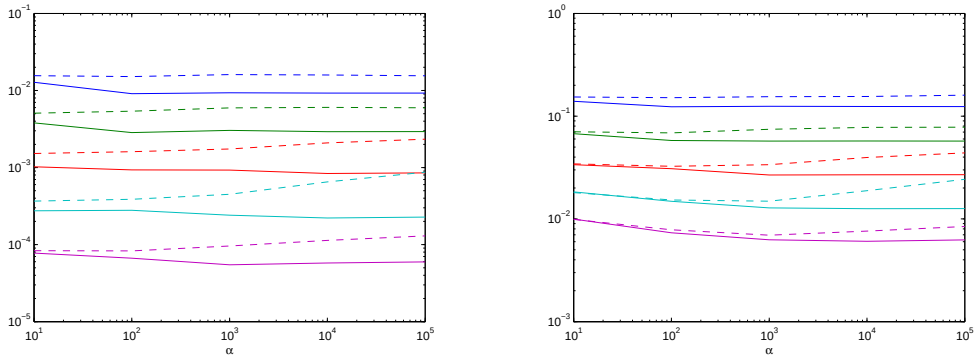
(a) L_2 norm error for Case II.

(b) H_1 semi-norm error for Case II.

Figure 2.10: Figures (a) and (b) show that the errors computed by our method (solid lines) are smaller than those by the standard finite element method (dash lines) among all $\widehat{\alpha}$ in Experiment 1 Case II. Each line represents the error versus $\widehat{\alpha}$ for fixed h. The values of h are 1/4, 1/8, 1/16, 1/32, 1/64 from top to bottom.

[68] applied on the coarse grid and we found that our method gives a consistently better performance for all values of $\widehat{\alpha}$, although the gain is less pronounced compared with that over the standard finite element method. From Fig. 2.11, we see that the errors computed by the IFE method increase as $\widehat{\alpha}$ increases. This shows that our method improves the accuracy and remove the dependence of $\widehat{\alpha}$.

Experiment 2. In this experiment Ω is the unit disk, α is as defined in (2.5.2),



(a) L_2 norm error for Case II.

(b) H_1 semi-norm error for Case II.

Figure 2.11: Figures (a) and (b) show that the errors computed by our method (solid lines) are smaller than those by the IFE method (dash lines) among all $\widehat{\alpha}$ in Experiment 1 Case II. Each line represents the error versus $\widehat{\alpha}$ for fixed h. The values of h are 1/4, 1/8, 1/16, 1/32, 1/64 from top to bottom.

with $r_0 = 1/3$, $f = 0$ and $g(x) = x$. The exact solution can be obtained analytically:

$$u(x,y) = \begin{cases} \frac{-2}{(\beta-1)\,r_0^2-(\beta+1)}x & , \quad r < 0, \\ \\ \frac{-(\beta+1)}{(\beta-1)\,r_0^2-(\beta+1)}x + \frac{(\beta-1)r_0^2}{(\beta-1)\,r_0^2-(\beta+1)}\frac{x}{x^2+y^2} & , \quad r \geq r_0, \end{cases} \tag{2.5.4}$$

where $\beta = \alpha_1/\alpha_0$. Unlike in Experiment 1, the exact solution depends on the polar angle θ. We investigate convergence for the case $\alpha_1 = \widehat{\alpha}$, $\alpha_0 = 1$, with increasing $\widehat{\alpha}$ (i.e., Case I), using quasi-uniform meshes, with a typical example shown in Figure 2.12. A typical numerical solution and pointwise error are shown in Fig. 2.13. As in Experiment 1, we see that the solution is flat in the high conductivity region and the errors are small and concentrated along the interface. From Table 2.3, we see the convergence rates are very close to optimal and independent of $\widehat{\alpha}$. Although the theory presented above is for polygonal $\partial\Omega$, the error estimates in Remark 2.4.19 could easily be extended to prove $\widehat{\alpha}$-independent convergence of optimal order for this experiment, as observed in the tables.

Experiment 3. In this experiment $\Omega = [-1,1] \times [-1,1]$ and we consider the case of two inclusions, with

$$\alpha = \begin{cases} \alpha_1\ , & \text{when} \quad (x-x_1)^2 + (y-y_1)^2 < r_1^2, \\ \alpha_2\ , & \text{when} \quad (x-x_2)^2 + (y-y_2)^2 < r_2^2, \\ \alpha_0\ , & \text{otherwise.} \end{cases}$$

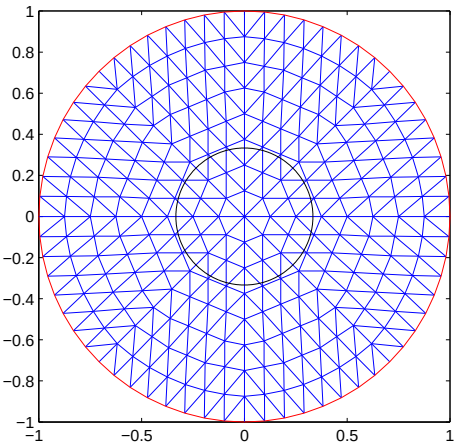
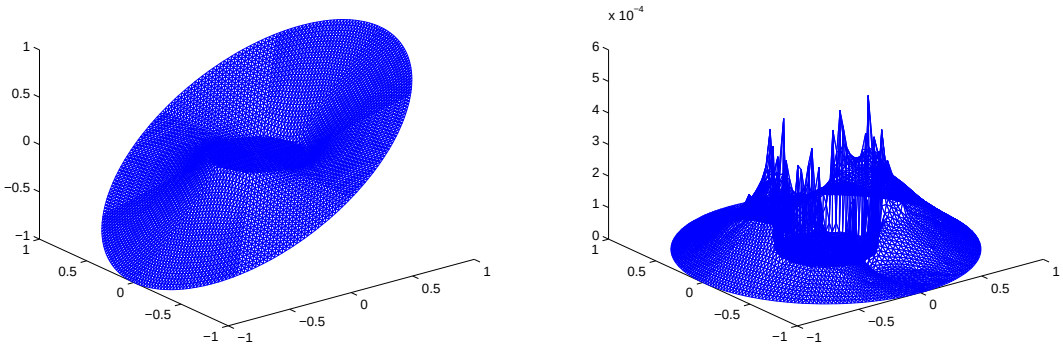


Figure 2.12: Illustration of computation domain and meshes for Experiment 2. The black circle indicates the interface in the problem.



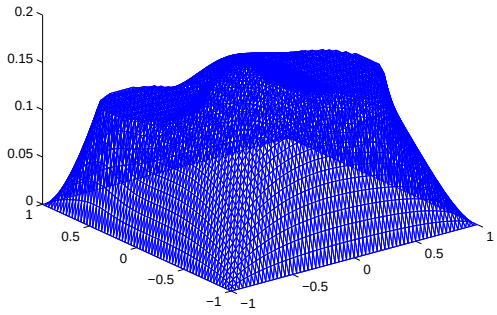
(a) u_h^{MS} (b) $|E_h^{\mathrm{MS}}|$

Figure 2.13: Numerical solution u_h^{MS} and pointwise error E_h^{MS} for Experiment 2 with $\widehat{\alpha} = 100000$ and $h = 1/32$.

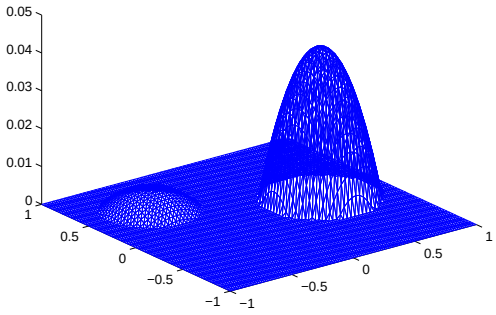
h	$\widehat{\alpha} = 10$	$\widehat{\alpha} = 100$	$\widehat{\alpha} = 1000$	$\widehat{\alpha} = 10000$	$\widehat{\alpha} = 100000$
1/8	2.2893e-3	4.6732e-3	3.4460e-3	3.3769e-3	3.3855e-3
1/16	7.0721e-4	1.7751e-3	9.0811e-4	8.8256e-4	8.8731e-4
1/32	1.8442e-4	3.1863e-4	2.5463e-4	2.4886e-4	2.8548e-4
1/64	5.2058e-5	7.9585e-5	7.0451e-5	7.0448e-5	7.0659e-5
rate	1.8315	2.0105	1.8671	1.8575	1.8383

h	$\widehat{\alpha} = 10$	$\widehat{\alpha} = 100$	$\widehat{\alpha} = 1000$	$\widehat{\alpha} = 10000$	$\widehat{\alpha} = 100000$
1/8	7.3816e-2	9.1749e-2	9.0765e-2	9.1247e-2	9.1505e-2
1/16	4.1501e-2	4.6103e-2	4.5586e-2	4.5827e-2	4.5973e-2
1/32	2.2267e-2	2.4132e-2	2.3906e-2	2.3967e-2	2.4874e-2
1/64	1.3250e-2	1.3547e-2	1.2411e-2	1.2333e-2	1.2382e-2
rate	0.8332	0.9213	0.9543	0.9597	0.9543

Table 2.3: The L_2-norm errors (upper) and the H^1 semi-norm errors (lower) for Experiment 2.

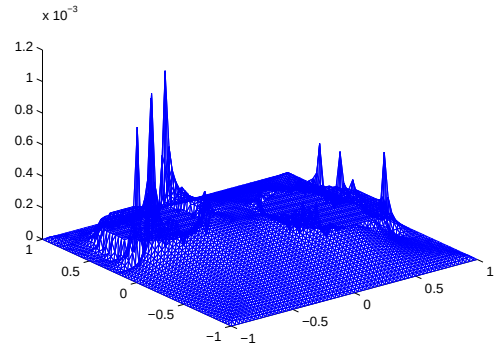


(a) u_h^{MS} for Case I with $\widehat{\alpha} = 100000$.

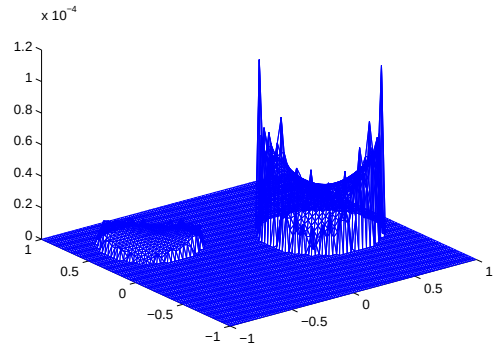


(b) u_h^{MS} for Case II with $\widehat{\alpha} = 100000$.

Figure 2.14: Numerical solutions u_h^{MS} for Experiment 3.



(a) $|E_h^{\mathrm{MS}}|$ for Case I with $\widehat{\alpha} = 100000$.



(b) $|E_h^{\mathrm{MS}}|$ for Case II with $\widehat{\alpha} = 100000$.

Figure 2.15: Pointwise errors E_h^{MS} for Experiment 3.

Here we choose $f = 1$ and $g = 0$. Since an analytical solution is unknown, we use the solution on the finest mesh (here with 16641 grid points) as the reference solution to compute the error for solutions on coarser meshes. We choose $(x_1, y_1) = (1/2, 0)$, $(x_2, y_2) = (-1/2, 1/2)$ and $r_1 = 2/5$, $r_2 = 1/3$.

Recalling that (2.1.5) allows α_1 and α_2 to approach infinity with different rates, we set $\alpha_1 = \widehat{\alpha}$, $\alpha_2 = 5\,\widehat{\alpha}$ and $\alpha_0 = 1$ as an example of Case I. For Case II, we let $\alpha_1 = 1$, $\alpha_2 = 5$ and $\alpha_0 = \widehat{\alpha}$. The numerical solution and pointwise error are shown in Fig. 2.14 and Fig. 2.15. We see that the error is still small and concentrated along the interface. Table 2.4 and 2.5 show that our method enjoys a roughly optimal convergence rate and the errors are independent of $\widehat{\alpha}$, as predicted by Theorem 2.4.18.

Discussion of the Choice of the Number of Subgrid Elements

h	$\widehat{\alpha} = 10$	$\widehat{\alpha} = 100$	$\widehat{\alpha} = 1000$	$\widehat{\alpha} = 10000$	$\widehat{\alpha} = 100000$
1/4	2.2150e-2	7.2008e-3	7.0956e-3	7.2309e-3	8.1280e-3
1/8	2.9498e-3	2.5855e-3	2.5863e-3	2.6387e-3	3.1500e-3
1/16	1.0142e-3	7.0168e-4	7.1187e-4	7.5761e-4	1.1063e-3
1/32	1.6523e-4	1.9426e-4	1.5366e-4	1.5484e-4	1.7776e-4
rate	2.2740	1.7518	1.8449	1.8436	1.8054

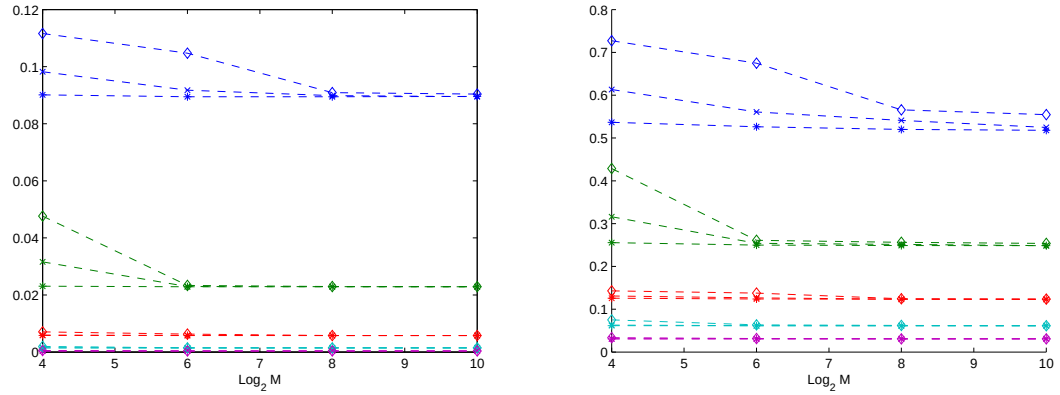
h	$\widehat{\alpha} = 10$	$\widehat{\alpha} = 100$	$\widehat{\alpha} = 1000$	$\widehat{\alpha} = 10000$	$\widehat{\alpha} = 100000$
1/4	1.0715e-1	3.5010e-2	3.5559e-2	3.5590e-2	3.5765e-2
1/8	2.2953e-2	1.5407e-2	1.5854e-2	1.6046e-2	1.8413e-2
1/16	1.2119e-2	5.9967e-3	6.7962e-3	7.9550e-3	1.6621e-2
1/32	5.8558e-3	3.1782e-3	2.0319e-3	2.2558e-3	4.0509e-3
rate	1.3502	1.1746	1.3610	1.2951	0.9574

Table 2.4: The L_2-norm errors (upper) and the H^1 semi-norm errors (lower) for the Case I: $\alpha_1 = \widehat{\alpha}$, $\alpha_2 = 5\,\widehat{\alpha}$, $\alpha_0 = 1$ in Experiment 3.

h	$\widehat{\alpha} = 10$	$\widehat{\alpha} = 100$	$\widehat{\alpha} = 1000$	$\widehat{\alpha} = 10000$	$\widehat{\alpha} = 100000$
1/4	6.6651e-3	1.3188e-3	1.2837e-3	1.2870e-3	1.2343e-3
1/8	7.7081e-4	4.9420e-4	3.3161e-4	3.3314e-4	3.2842e-4
1/16	1.8952e-4	1.6737e-4	6.9212e-5	6.9211e-5	6.6224e-5
1/32	4.9532e-5	5.8600e-5	1.4979e-5	1.5999e-5	1.6192e-5
rate	2.3240	1.5039	2.1524	2.1257	2.1067

h	$\widehat{\alpha} = 10$	$\widehat{\alpha} = 100$	$\widehat{\alpha} = 1000$	$\widehat{\alpha} = 10000$	$\widehat{\alpha} = 100000$
1/4	5.9426e-2	9.7096e-3	8.7966e-3	8.7811e-3	8.1776e-3
1/8	1.1250e-2	6.9732e-3	4.4793e-3	4.2302e-3	4.4192e-3
1/16	4.7793e-3	5.0584e-3	1.3624e-3	1.2180e-3	1.1965e-3
1/32	3.0188e-3	2.9506e-3	6.1242e-4	4.0833e-4	4.7003e-4
rate	1.4132	0.5618	1.3250	1.5076	1.4248

Table 2.5: The L_2-norm errors (upper) and the H^1 semi-norm errors (lower) for the Case II: $\alpha_1 = 1$, $\alpha_2 = 5$, $\alpha_0 = \widehat{\alpha}$ in Experiment 3.



(a) L_2 norm error versus M

(b) H^1 semi-norm error versus M.

Figure 2.16: The sensitivity test of M for Case I in Experiment 1 with $\widehat{\alpha} = 100$(- - * line), 1000(- - x line), 10000(- - diamond line).The values of h are 1/4, 1/8, 1/16, 1/32, 1/64 from top to bottom.

Finally we discuss the sensitivity of the error in the overall multiscale method to the choice of M (the number of subgrid elements) used to compute the basis functions. Here we test how the errors depend on M in Experiment 1 (Case I), with $\widehat{\alpha} = 100, 1000, 10000$, and $h = 1/4, 1/8, 1/16, 1/32, 1/64$ and plot the error against M in Fig. 2.16. In each plot there are five groups of three lines; each group corresponds to a different value of h, with h decreasing from 1/4 (top group) to 1/64 (bottom group). In each group of three lines, the dash-asterisk line is for $\widehat{\alpha} = 100$, the dash-x line is for $\widehat{\alpha} = 1000$ and the dash-diamond line is for $\widehat{\alpha} = 10000$. (In the case of the L_2 error, the cases $h = 1/32$ and 1/64 are almost coincident.)

From these graphs we see that the errors decrease as M increases, and with $M = 64$, the multiscale finite element method gives an error that is comparable to that using $M = 1024$ when the coarse mesh size is less then 1/4 for all $\widehat{\alpha}$. This indicates that, at least in this example, it is possible to use relatively few subgrid elements to compute the basis function with the desired accuracy, for example by choosing $M = 64$. We expect that the use of adaptive subgrid elements may lead to further computational savings in computing the multiscale basis function.

2.6 Concluding Remarks and Discussions

In this work we propose and analyze a new multiscale finite element method for the elliptic interface problem. From Theorem 2.2.2, we see that the boundary conditions of basis functions play an essential role in error control. The most important con-

tribution of this work is finding the boundary conditions for the basis functions to obtain optimal convergence rates independent of the contrast. We provide the complete analysis for error estimates without scale separation assumptions. This has not been done by previous studies in the MsFEM framework.

Our method can be considered as a generalization of the immersed finite element method (IFE) proposed by Li, Lin and Wu [68]. Our method improves IFE in several aspects. The analysis of the IFE method relies on strong regularity assumptions of the exact solution u, which has not been proved and may not be true in general. Specifically, Li, Lin and Wu assumed that u is C^2 in each sub-regions, and the constants in their error estimates for the IFE method depend on the contrast $\widehat{\alpha}$ and the C^2-norm of u. On the other hand, our analysis requires only $f \in H^{1/2}(\Omega)$. The numerical experiments in Section 2.5 indicate that the convergence of IFE has some mild dependence on the contrast, $\widehat{\alpha}$, while the convergence of our multiscale finite element method is completely independent of the contrast. However, the computational cost of our method is slightly higher than that of the IFE method since we need to construct the basis functions numerically.

The standard finite element method with body fitted mesh for interface problem has optimal convergence rates and the hidden constants are independent of the contrast [82]. With today's well-developed mesh generating software and fast linear solvers, the standard finite element method with body fitted mesh could be more efficient and accurate than our method. On the other hand, the multiscale finite element approach breaks down a large system into many small local problems. The basis functions can be computed locally and can be carried out in perfect parallel. Moreover, the MsFEM approach reduces the size of the computation which offers a big saving in computer memory. This is useful when dealing with complicated domains. For example, Strouboulisl, Zhang and Babuška [86] considered the Laplace equation in a square domain minus 597 voids. To resolve the voids, it requires 30372 quadrilateral elements with 133180 degrees of freedom (See Fig. 2.6). Strouboulisl et al used generalized finite element method on the 8x8 mesh with 243 degrees of freedom to solve the problem and obtained a satisfactory accuracy. Although our method cannot be applied to such a complicated problem yet, it may be possible to extend our method to this problem in the future.

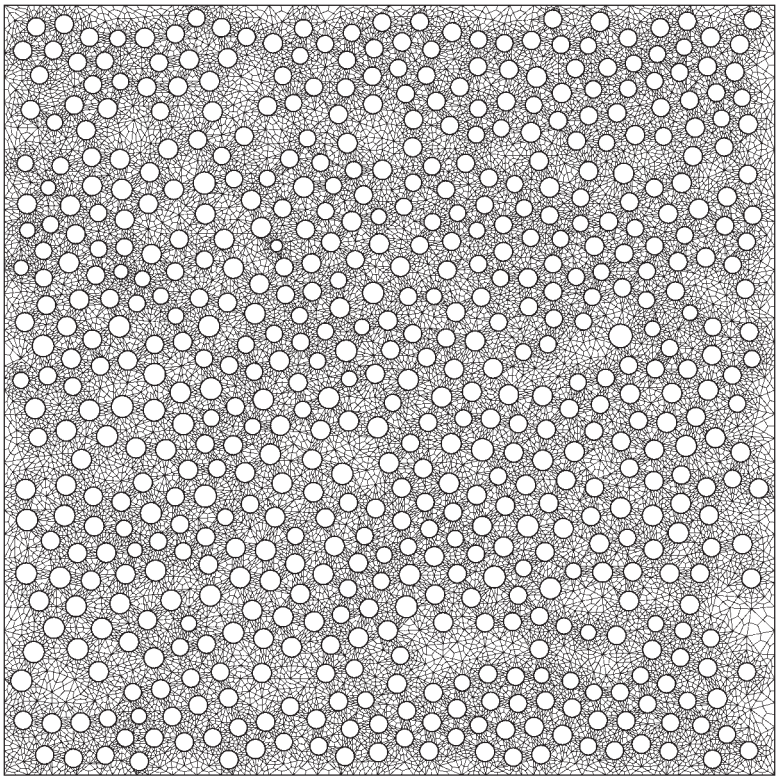


Figure 2.17: An FEM mesh with 30372 quadrilateral elements used. This figure is from [86].

Chapter 3

Flow-Based Oversampling Techniques for Two-Phase Flow Problem

The work presented in this chapter consists of materials from the paper [30] by Chu et al. In the next section we give a preliminary explanation of the multiscale finite volume element method (MsFVEM). In Section 3.2, we present the flow-based oversampling approach and analysis. Finally, in Section 3.3, numerical results are presented.

3.1 The Multiscale Finite Volume Element Procedure

Mass conservative schemes play an important role in groundwater simulation. Therefore, it is essential to use methods which provide a mass conservative approximation to the flux $\mathbf{v} = -k\nabla p$. In this section, we briefly recall the multiscale finite volume element method (MsFVEM) [55, 38], one of the popular mass conservative methods.

To demonstrate the concept of MsFVEM, we denote by $\mathcal{K}^h$ the set of coarse elements (rectangles in this case) K and $Z_h(K)$ the set of the vertices of K. The quantity ξ_K indicates the center of coarse element K. Element K is divided into four rectangles of equal area by connecting ξ_K to the midpoints of the element edges. These quadrilaterals are denoted by K_ξ, where $\xi \in Z_h(K)$, are the vertices of K. We designate $Z_h = \bigcup_K Z_h(K)$ and $Z_h^0 \subset Z_h$ the vertices which do not lie on the Dirichlet boundary of Ω. The control volume V_ξ is defined as the union of the quadrilaterals K_ξ sharing the vertex ξ. The grid comprised of elements K (solid squares in Figure 3.1) is sometimes referred to as the primal grid and the grid defined by V_ξ (dashed

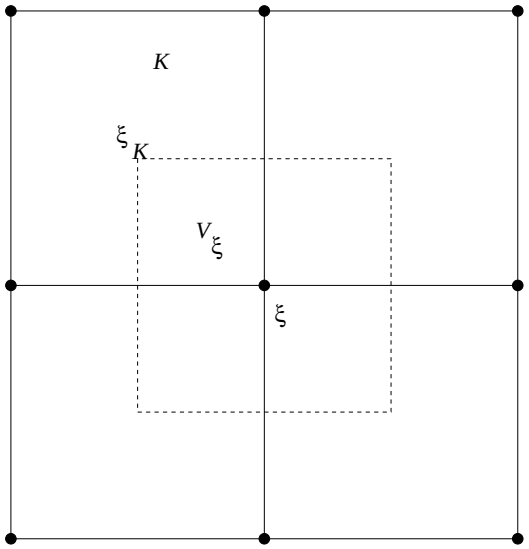


Figure 3.1: Schematic of nodal points and grid.

square in Figure 3.1) as the dual grid. In our procedure we compute pressure at the vertices of the primal grid. This differs from the approach of [55, 56] in which pressure is computed at the centroids of the primal grid blocks. This also leads to a different treatment of global boundary conditions.

The goal of the MsFVEM is to determine coarse scale basis functions that incorporate the fine scale information in the underlying permeability description. The technique applied here follows the multiscale finite element method of [51], as the basis functions are determined from the solution of the leading order homogeneous elliptic equation on each coarse element. For a coarse rectangular element K, the basis functions $\phi_i,\ i = 1, 2, 3, 4$, are computed via solution of:

$$\begin{aligned} \nabla \cdot (\mathbf{k} \cdot \nabla \phi_i) &= 0 \quad \text{in } K \\ \phi_i &= g_i \quad \text{on } \partial K, \end{aligned} \tag{3.1.1}$$

for prescribed boundary functions g_i. The basis function associated with the vertex $\mathbf{x}_i$ is constructed from the union of the basis functions that share this $\mathbf{x}_i$ and are zero elsewhere. Note that ϕ_i must satisfy $\phi_i(\mathbf{x}_j) = \delta_{ij}$.

Hou and Wu [51] showed that the accuracy of the resulting coarse model is impacted by the treatment of boundary effects in (3.1.1). Enhanced accuracy can be achieved by solving local 1D problems [55] for the determination of g_i or, as is considered here, by solving (3.1.1) in a domain that includes more than just the fine scale cells corresponding to the coarse block K (this approach is referred to as oversam-

pling). The specific boundary conditions that are used in this paper for the determination of the basis functions will be discussed in detail below. A vertex-centered finite volume procedure is used to solve (3.1.1).

As discussed in [37], once the basis functions are constructed we determine $p^h \in V^h$, where V^h is the space of approximate pressure solutions, with $p^h = \sum_{\mathbf{x}_j \in Z_h^0} p_j \phi_j$, by enforcing

$$\int_{\partial V_\xi} \left(\lambda(S)\mathbf{k} \cdot \nabla p^h\right) \cdot \mathbf{n}\, dl = \int_{V_\xi} q d\mathbf{x}, \tag{3.1.2}$$

for every control volume $V_\xi \subset \Omega$. Here $\mathbf{n}$ defines the normal vector on the boundary of the control volume ∂V_ξ and S is the fine scale saturation. Note that the integral in (3.1.2) is performed over a coarse cell in the dual grid (V_ξ) and the finite element test function is unity. For this reason, the technique is referred to as a finite volume element method. In this way the method differs from multiscale finite element procedures (e.g., [51]). The equation (3.1.2) results into a system of linear equations for the solution values at the nodal points of the coarse mesh. In particular, we have

$$A\mathbf{p} = \mathbf{b},$$

where $A = (a_{ij})$, $a_{ij} = \int_{\partial V_{\xi_i}} (\lambda(S)\mathbf{k} \cdot \nabla \phi_j) \cdot \mathbf{n}\, dl$, $b_i = \int_{V_{\xi_i}} q d\mathbf{x}$.

3.2 Flow-Based Oversampling for Multiscale Finite Element Methods

First, we describe the oversampling technique. Denote a target coarse block by K (we assume rectangular partition in 2D, for simplicity) and an extended coarse region by K' (see Figure 3.2). For K' with vertices $\mathbf{y}_i$ ($i = 1, 2, 3, 4$), we denote by $\psi_i(\mathbf{x})$ a nodal basis on K', such that $\psi_i(\mathbf{y}_j) = \delta_{ij}$. The nodal basis functions ψ_i ($i = 1, 2, 3, 4$) are constructed by solving (3.1.1) in the region K' (see Figure 3.2) with *linear boundary conditions*. Once the auxiliary functions ψ_i (also called oversampling functions) are constructed, we compute the basis functions ϕ_i as a linear combination of the ψ_i (as

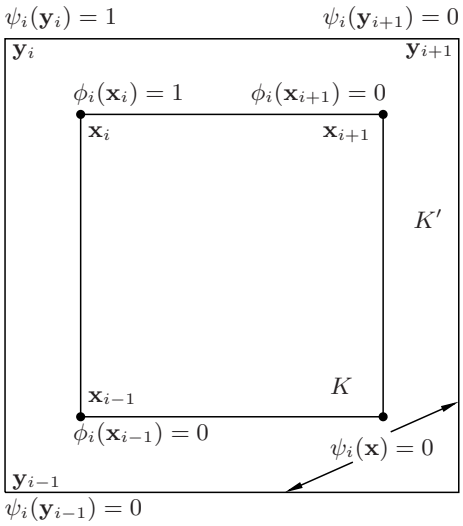


Figure 3.2: Schematic description of coarse block and extended coarse block regions.

is done in oversampling for MsFEM [51]) as follows:

$$\phi_i(\mathbf{x}) = \sum_{j=1}^{4} c_{ij}\psi_j(\mathbf{x}), \tag{3.2.1}$$

where $\mathbf{x}_j$ are the nodes of the target coarse block K and c_{ij} are coefficients determined by imposing $\phi_i(\mathbf{x}_j) = \delta_{ij}$. The resulting multiscale basis functions are nonconforming. Using these basis functions, the global problem is solved using (3.1.2). We emphasize that this method is not limited to rectangular global domains. In the case of non-rectangular domains, one can still use global auxiliary solutions with some generic boundary conditions. These auxiliary fields are required to be linearly independent, so that one can construct linearly independent multiscale basis functions ϕ_i.

In our simulations, we are interested in taking the oversampled domain to be the entire region, i.e., $K' = \Omega$. This is in particular used for porous media without scale separation which exhibit strong non-local effects, such as those considered in this work. By taking the oversampling region to be the entire domain, one avoids the resonance errors endemic to numerical homogenization (see, e.g., [51, 41]). The ψ_i computed in this manner are the global solutions corresponding to single-phase flow problems which are computed once. We refer to the *standard oversampling* technique when generic oversampling functions are used (as described above) for constructing multiscale basis functions.

In this work, we propose the use of flow-based oversampling auxiliary functions

for the construction of basis functions which differs from the standard oversampling method. In the flow-based oversampling method, we replace some of the standard oversampling auxiliary functions (solutions of single-phase flow equations) by ones obtained from solving the single-phase flow problem with the actual boundary conditions of two-phase flow. More precisely, if the two-phase flow equations are solved subject to some boundary conditions, then we replace some of the generic oversampling auxiliary functions, ψ_j in (3.2.1), with oversampling functions with the same boundary conditions as those in the two-phase flow. Note that the standard oversampling assumes that the oversampling functions have linear boundary conditions. In applications, the boundary conditions are often non-smooth and can have an impact on the flow solutions. For this purpose, one needs to take into account non-smooth effects via auxiliary oversampling functions. Such situations occur, for example, when the flow is corner-to-corner (see the next section for details). In this case, the boundary conditions are no longer smooth and one of the oversampling functions is taken to be the solution of corner-to-corner flow.

We remark that the computational time required for the standard oversampling and the flow-based oversampling is similar. Indeed, in the flow-based oversampling approach, some of the standard oversampling auxiliary functions are replaced by flow-based single-phase flow solutions. The latter does not affect the computational time unless we incorporate many flow boundary conditions into the multiscale basis functions. However, the computational time required for flow-based approaches is, generally, larger than that required for local multiscale methods. Since in local approaches one still needs to solve the local problems over each coarse grid block, the total computational time can be similar to solving the global problems. This holds particularly if the oversampling regions are larger than the target coarse grid blocks. We stress again that the global solutions are computed off-line for the calculation of the multiscale basis functions. One can incorporate multiple sources of global information in the flow-based oversampling approach, in the form of multiple multiscale bases, which is not the case with the multiscale finite element method introduced in [37, 35] which uses limited global information.

3.2.1 Analysis

In this section, we show that the use of the actual flow boundary conditions in oversampling methods is important. For flow based oversampling approach, we can see that the solution of the single-phase flow problem with exact boundary conditions lies in the approximation space V^h. Solutions with generic boundary conditions also lie in the approximation space V^h. Our analysis shows that the two-phase flow solution can be well approximated in V^h if we add flow-based oversampling basis.

Consider the flow equations for two-phase flow in the form

$$-\nabla \cdot (\lambda(S)\mathbf{k} \cdot \nabla p) = 0,\ p|_{\Gamma_1} = g_D,\ \tfrac{\partial p}{\partial n}|_{\Gamma_2} = g_N; \tag{3.2.2}$$

where g_D and g_N may be discontinuous along the boundaries Γ_1, Γ_2. Let Q be the solution for the single phase flow equation,

$$-\nabla \cdot (\mathbf{k} \cdot \nabla Q) = 0,\ Q|_{\Gamma_1} = g_D,\ \tfrac{\partial Q}{\partial n}|_{\Gamma_2} = g_N. \tag{3.2.3}$$

Using equation (3.2.2), and (1.4.3) (with $h_w = 0$), we derive the equation for p:

$$\begin{aligned} -\nabla \cdot (\mathbf{k} \cdot \nabla p) &= -\frac{1}{\lambda(S)} \nabla \cdot (\lambda(S)\mathbf{k} \cdot \nabla p) - \frac{1}{\lambda(S)^2}(\lambda(S)\mathbf{k} \cdot \nabla p) \cdot \lambda'(S)\nabla(S) \\ &= -\frac{\lambda'(S)}{\lambda(S)^2}\mathbf{v} \cdot \nabla(S) \\ &= \frac{1}{f'(S)} \left(\frac{1}{\lambda(S)}\right)_t \end{aligned} \tag{3.2.4}$$

with $p|_{\Gamma_1} = g_D,\ \frac{\partial p}{\partial n}|_{\Gamma_2} = g_N$. Now let $w = p - Q$; subtracting equations (3.2.2) and (3.2.4) we get

$$-\nabla \cdot (\mathbf{k} \cdot \nabla w) = \tfrac{1}{f'(S)} \left(\tfrac{1}{\lambda(S)}\right)_t,\ w|_{\Gamma_1} = 0,\ \tfrac{\partial w}{\partial n}|_{\Gamma_2} = 0. \tag{3.2.5}$$

Thus the difference between p and Q satisfies the elliptic equation with a nonzero source term and homogeneous boundary conditions. Therefore, if $\frac{1}{f'(S)}(\frac{1}{\lambda(S)})_t$ is sufficiently regular, w can be approximated by generic solutions very well. Thus $p = Q + w$, where w is a smooth function of two linearly independent solutions of single-phase flow equation, as was shown in [78]. This result further justifies the use of Q, the solution of the single-phase flow problem with the actual flow boundary conditions.

3.3 Numerical Results

In this section, we present representative simulation results with quadratic relative permeabilities, $k_{rw}(S) = S^2$ and $k_{ro}(S) = (1-S)^2$. In all cases the system uses permeability fields from one of the upper Ness layers in the SPE comparative project [28], a benchmark test. These permeability fields are highly heterogeneous, channelized, and difficult to upscale. Figure 3.3 depicts the log-permeability of one of the layers.

We employ the flow-based multiscale finite volume element method for solving the pressure equation (1.4.2). The basis functions are constructed once and used throughout the simulations without updating them. The saturation equation (1.4.3) is solved on the fine grid using upwind finite volume method with flux limiter. For this purpose, the fine-scale velocity field is reconstructed from the multiscale basis functions representation of the pressure via Darcy's Law.

Simulation results are presented for the saturation fields, as well as total flow rate and the oil cut as a function of pore volume injected (PVI). The oil cut is also referred to as the fractional flow of oil. The oil cut (or fractional flow) defined as the fraction of oil in the produced fluid, is given by q_o/q_t, where $q_t = q_o + q_w$. Here q_o and q_w are the flow rates of oil and water at the production edge of the model. In particular, $q_w = \int_{\partial\Omega^{out}} f(S)\mathbf{v}\cdot\mathbf{n}d\omega$, $q_t = \int_{\partial\Omega^{out}} \mathbf{v}\cdot\mathbf{n}d\omega$, and $q_o = q_t - q_w$, where $\partial\Omega^{out}$ is the outer flow boundary. We use the notation Q for total flow q_t and F for fractional flow q_o/q_t in numerical results. Pore volume injected, defined as PVI $= \frac{1}{V_p}\int_0^t q_t(\tau)d\tau$, with V_p being the total pore volume of the system, provides the dimensionless time for the displacement. When using multiscale finite volume element methods for two-phase flow, one can update the basis functions near the sharp fronts. Indeed, sharp fronts modify the local heterogeneities and this can be taken into account by resolving the local equations, (3.1.1), for basis functions. If the saturation is smooth in the coarse block, it can be approximated by its average in (3.1.1), and consequently, the basis functions do not need to be updated. It can be shown that this approximation yields first-order errors (in terms of coarse mesh size). In our simulations, we found only slight improvement when the basis functions are updated, thus the numerical results for the MsFVEM presented here do not include the basis functions update near the sharp fronts.

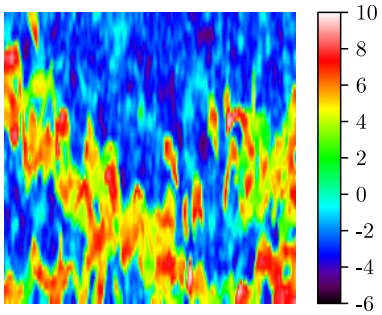


Figure 3.3: Log-permeability for one of the layers of upper Ness.

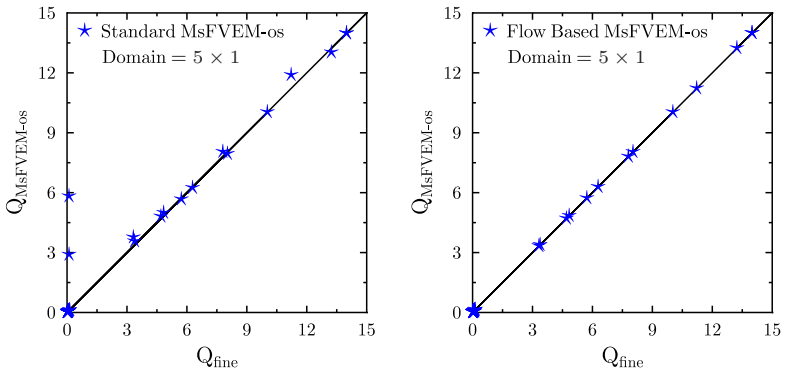


Figure 3.4: Total single-phase flow rate for 50 layers of SPE 10 using the standard oversampling (left) and the flow-based oversampling (right). Domain size is 5×1.

In all numerical examples, the fine-scale field is 220×60, while the coarse-scale field is 22×6. We have observed similar results for other coarse grids. The boundary condition is imposed by specifying $p = 1$, $S = 1$ along the $x = 0$ edge for $0 \leq z \leq 0.1$ and $p = 0$ along the $x = L$ edge for $0.9 \leq z \leq 1$. On the remainder of the boundary, we assume a no flow boundary condition. Here, L is the horizontal size of the global rectangular domain. We note that these boundary conditions are different from those used in constructing generic oversampling functions.

In our first numerical results, we compare the total flow rate Q for single-phase flow using the standard oversampling and the flow-based oversampling methods. The global domain sizes are 5×1 (Figure 3.4) and 20×1 (Figure 3.6), i.e., $L = 5$ and $L = 20$, respectively. In both cases, we observe that the flow-based oversampling gives nearly exact results, while standard oversampling methods are not as accurate. The accuracy of standard oversampling methods deteriorates as the anisotropy ratio increases.

Next, we present numerical results for dynamic quantities, such as fractional flow, total flow rate and saturation maps for two-phase flow and transport. In Figures

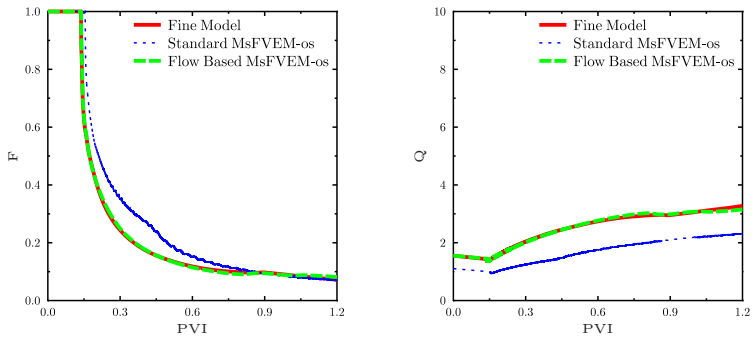


Figure 3.5: Fractional flow (left figure) and total production (right figure) comparison for the standard oversampling and the flow-based oversampling techniques. The viscosity ratio is $\mu_o/\mu_w = 5$.

3.5 and 3.7, we present the fractional flow ($F = q_o/q_t$, left figure) and the total flow ($Q = q_t$, right figure) for two viscosity ratio cases, $\mu_o/\mu_w = 5$ and $\mu_o/\mu_w = 15$. The solid line designates the fine-scale reference solution, while the dotted line designates the standard oversampling method where generic global single-phase flow solutions are used, and the dashed line designates the flow-based oversampling method. We observe from these figures that the flow-based oversampling method is more accurate. This is more evident from the total flow plot (Figure 3.4 and 3.6). Next, we compare the saturation fields at different time instances. In Figure 3.8, the saturation fields at the time instances, PVI $= 0.1, ..., 0.9$ are depicted. One observes that the saturation fields obtained from the standard oversampling method are not very accurate. This is more evident in the regions close to the upper right corner.

To compare the saturation maps at different time instances quantitatively, we plot L_2 errors of the saturations fields in Figure 3.9. In particular, we present the errors for 4 different layers of SPE 10 with two different viscosity ratios $\mu_0/\mu_w = 5$ (left figures) and $\mu_o/\mu_w = 15$ (right figures). The saturation errors are computed at $PVI = 0.1, 0.2, ..., 0.9$. We observe that the errors in saturation fields from flow-based oversampling techniques are smaller compared to these from the standard oversampling method in all cases. In most cases, the error is smaller by several factors. We note again that the cost of these computations are the same and involve computing four auxiliary oversampling functions. We have also tried numerical results with different boundary conditions where Dirichlet boundary conditions are imposed over two coarse grid blocks at $x = 0$ and $x = L$. In these cases, we have observed consistent improvement when the flow-based oversampling method is used.

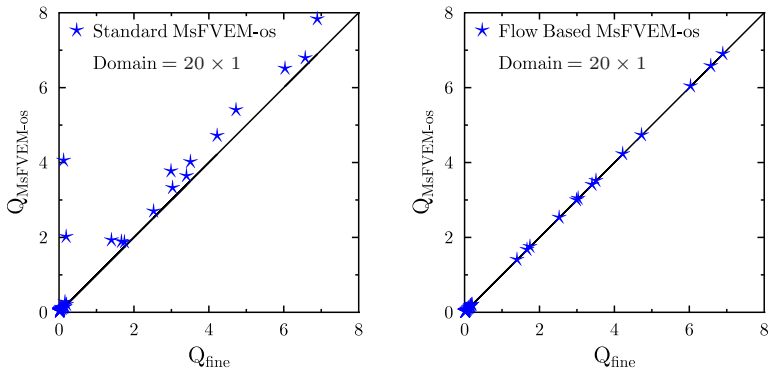


Figure 3.6: Total single-phase flow rate for 50 layers of SPE 10 using the standard oversampling (left) and the flow-based oversampling (right). Domain size is 20×1.

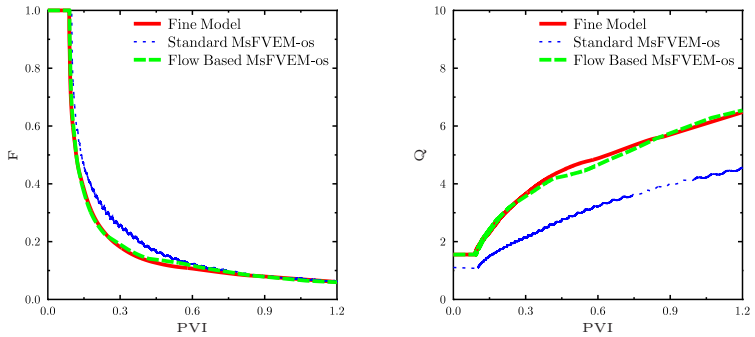


Figure 3.7: Fractional flow (left figure) and total production (right figure) comparison for the standard oversampling and the flow-based oversampling techniques. The viscosity ratio is $\mu_o/\mu_w = 15$.

In the work, we discussed the cases with singular type boundary conditions and assumed no source terms. Source terms representing well information are common in applications. One can consider singular source terms (such as Dirac δ functions) within the proposed framework. In this case, the flow-based auxiliary functions will be the global solutions containing the singular sources. We will present these results in a future paper.

3.4 Concluding Remarks

In this work, we study oversampling techniques for multiscale simulation of two-phase immiscible flow in heterogeneous porous media with strong non-local effects. We propose the flow-based oversampling technique where the actual two-phase flow boundary conditions are used to construct the oversampling functions. In our numerical simulations, the oversampling region is taken to be the entire domain to capture strong

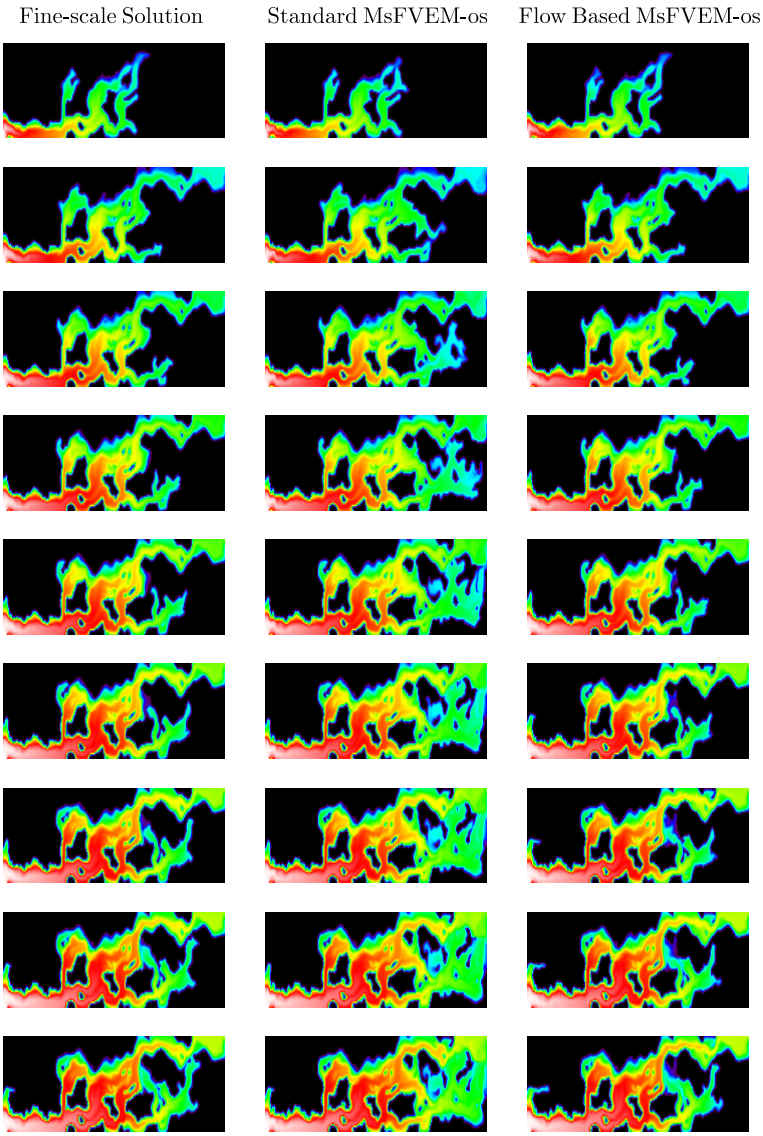


Figure 3.8: Saturation maps at PVI=0.1–0.9 for fine-scale solution (left figure), standard MsFVEM-os (middle figure), and flow-based MsFVEM-os (right figure). Corner-to-corner boundary condition is used. Layer 61 in SPE 10.

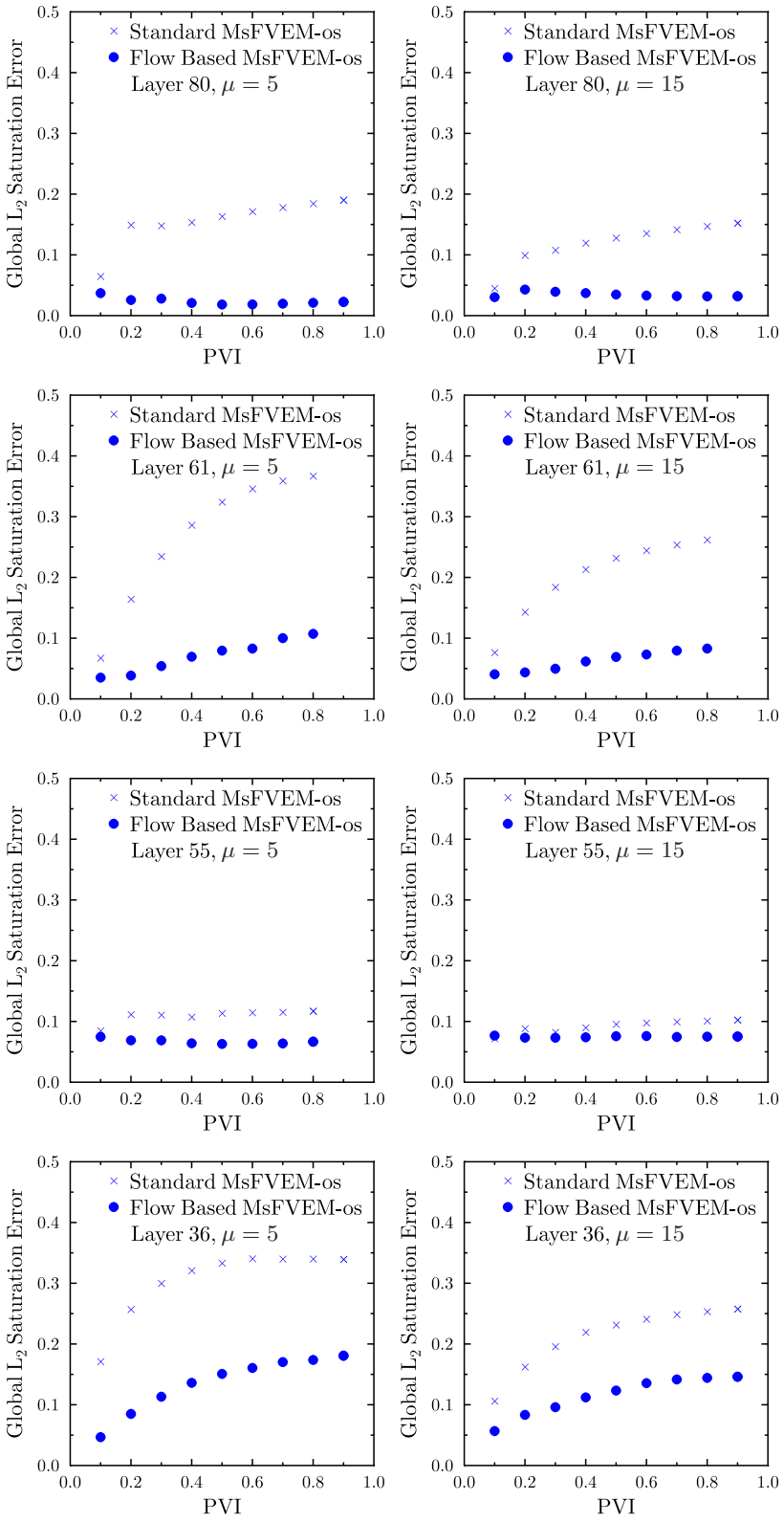


Figure 3.9: L_2 saturation errors at different PVIs.

non-local effects. We compare our approach to standard oversampling technique, which uses generic global boundary conditions that do not reflect the actual flow boundary conditions. The flow-based oversampling approach replaces some of the global oversampling basis functions with the solutions of single-phase flow equation with actual two-phase flow boundary conditions. Our numerical results show that the second approach is several times more accurate in almost all the cases considered. We provide a partial theoretical explanation of these numerical observations.

Chapter 4

Metric Based Upscaling Methods for the Pressure Equation

Many multiscale methods [27, 15, 34] for two-phase flow systems have been proposed over the decades. However, most of them can be only demonstrated numerically. A complete analysis of the pressure and saturation is very difficult. In this chapter, we focus on the pressure equations in two-phase flow problems

$$\begin{aligned} -\nabla \cdot (\lambda(\mathbf{x})\mathbf{k}(\mathbf{x}) \cdot \nabla p(\mathbf{x})) &= f \quad \text{in} \quad \Omega, \\ p &= 0 \quad \text{on} \quad \partial\Omega, \end{aligned} \tag{4.0.1}$$

where Ω is a bounded convex domain, $\mathbf{k}$ is a $n \times n$ symmetric positive definite matrix with entries in $L_\infty(\Omega)$ and f is a $L_\infty(\Omega)$ function. We generalize the metric based upscaling techniques introduced by Owhadi and Zhang [78] to solve (4.0.1) when the mobility function λ is a positive $C^1(\bar{\Omega})$-function.

4.1 Compensation Phenomena

We begin with an introduction to the so-called $\mathbf{k}$-harmonic coordinates $F(\mathbf{x})$ associated with

$$\begin{aligned} -\nabla \cdot (\mathbf{k}\nabla p) &= f \quad \text{in} \quad \Omega, \\ p &= 0 \quad \text{on} \quad \partial\Omega. \end{aligned} \tag{4.1.1}$$

The harmonic coordinates F are defined by $F(\mathbf{x}) = (F_1(\mathbf{x}), ..., F_n(\mathbf{x}))$, where each component F_i is the solution of the homogeneous partial differential equation

$$\begin{aligned} \nabla \cdot (\mathbf{k}\nabla F_i) = 0 \quad &\text{in} \quad \Omega, \\ F_i(\mathbf{x}) = x_i \quad &\text{on} \quad \partial\Omega. \end{aligned} \tag{4.1.2}$$

Alessandrini and Nesi [5, 6, 7] studied the properties of the **k**-harmonic coordinates due to its origin in application to homogenization. They proved that F is always a homeomorphism in two dimensions [6, 7] and connected their result to homogenization when **k** has periodic structures [6].

Owhadi and Zhang [78] showed that the **k**-harmonic coordinates also play an important role in homogenization for non-periodic **k**. They discovered that the solution p of (4.1.1) is $W^{2,r}$ with respect to the harmonic coordinates, although it is only $W^{1,r}$ in Euclidean coordinates. When F is an automorphism, they converted (4.1.1) to the harmonic coordinates $\mathbf{y} = F(\mathbf{x})$ and obtained the elliptic equation in non-divergence form:

$$\begin{aligned} -\sum_{i,j=1}^{n} Q_{ij} \frac{\partial}{\partial y_i}\frac{\partial}{\partial y_j} w = \tilde{f} \quad &\text{in} \quad \Omega, \\ w = 0 \quad &\text{on} \quad \partial\Omega, \end{aligned} \tag{4.1.3}$$

where

$$w(\mathbf{y}) = p \circ F^{-1}(\mathbf{y}), \quad \tilde{f}(\mathbf{y}) = \left(\frac{f}{|\det(\nabla F)|}\right) \circ F^{-1}(\mathbf{y}), \tag{4.1.4}$$

and

$$Q(\mathbf{y}) = \left(\frac{(\nabla F)^t \mathbf{k} \nabla F}{|\det(\nabla F)|}\right) \circ F^{-1}(\mathbf{y}). \tag{4.1.5}$$

Maugeri, Palagachev and Softova proved that the solution u of the linear Dirichlet problem

$$\begin{aligned} \sum_{i,j=1}^{n} a_{ij} D_{ij} u = f \quad &\text{in} \quad \Omega, \\ u = 0 \quad &\text{on} \quad \partial\Omega, \end{aligned} \tag{4.1.6}$$

belongs to $W_0^{2,r}(\Omega)$ for some $r > 2$, if the uniformly elliptic coefficient a_{ij} satisfies the

Cordes condition [74]. The Cordes condition takes the form

$$\frac{\sum_{i,j=1}^{n} a_{ij}(x)^2}{[\sum_{i=1}^{n} a_{ii}(x)]^2} \leq \frac{1}{n-1+\varepsilon}, \quad \text{a.e} \quad \Omega \tag{4.1.7}$$

for some $0 < \varepsilon < 1$. If in addition a_{ij} is a symmetric matrix, (4.1.7) can be written as

$$\frac{\sum_{i=1}^{n} \lambda_i(a(x))^2}{[\sum_{i=1}^{n} \lambda_i(a(x))]^2} \leq \frac{1}{n-1+\varepsilon}, \quad \text{a.e} \quad \Omega, \tag{4.1.8}$$

where $\lambda_i(a(x))$ is the i-th eigenvalue of $a(x)$.

Owhadi and Zhang [78] applied the same techniques to analyze (4.1.3). They defined σ by

$$\sigma = (\nabla F)^t \mathbf{k} \nabla F,$$

and β_σ the Cordes parameter associated to σ by

$$\beta_\sigma = \operatorname{esssup}_{x\in\Omega} \left(n - \frac{(\operatorname{Trace}[\sigma])^2}{\operatorname{Trace}[\sigma^t \sigma]} \right).$$

The coefficient σ is called stable if and only if $\beta_\sigma < 1$ and $\|(\operatorname{Trace}[\sigma])^{n/2r-1}\|_{L_\infty(\Omega)} < \infty$. The definition is a modification of the Cordes condition for equation (4.1.3). Define the notation $\lambda_{\min}(M) = \operatorname{essinf}_{\mathbf{x}\in\Omega} \lambda_{\min}(M(\mathbf{x}))$ and $\lambda_{\max}(M) = \operatorname{essup}_{\mathbf{x}\in\Omega} \lambda_{\max}(M(\mathbf{x}))$ for matrix-valued functions M on Ω. Using the equivalent relation between (4.1.1) and (4.1.3), and a regularity theorem analogous to Theorem 1.2.3 of [74], Owhadi and Zhang proved the following theorem.

Theorem 4.1.1. *(Owhadi and Zhang [78]) Assume that σ is stable and F is an automorphism on Ω, then there exist constants $r > 2$ and $C > 0$ such that $p \circ F^{-1} \in W_0^{2,r}(\Omega)$ and*

$$\|p \circ F^{-1}\|_{W_0^{2,r}(\Omega)} \leq C \|f\|_{L_r(\Omega)} \tag{4.1.9}$$

The constant r depends on n, Ω, $\lambda_{\max}(\mathbf{k})$, $\lambda_{\min}(\mathbf{k})$, and β_σ. The constant C depends on the constants above and $\|(Trace(\sigma))^{\frac{n}{2r}-1}\|_{L^r(\Omega)}$.

Remark 4.1.2. In 2D, the $\mathbf{k}$-harmonic coordinates F is always an automorphism [7] and the Cordes condition holds trivially. Therefore, the assumption of Theorem 4.1.1 can be simplified to

$$0 < \lambda_{\min}(\sigma) \leq \lambda_{\max}(\sigma) < \infty,$$

which can be checked more easily. Moreover, one can show that there exists constants $\alpha > 0$ and r such that $p \circ F^{-1} \in C^{1,\alpha}(\Omega)$ and

$$\|\nabla(p \circ F^{-1})\|_{C^{\alpha}(\Omega)} \leq C\|f\|_{L_r(\Omega)}, \tag{4.1.10}$$

by the Sobolev embedding inequality and Theorem 4.1.1.

4.2 Numerical Homogenization for the Pressure Equation

Now we come back to the pressure equation:

$$\begin{aligned} -\nabla \cdot (\lambda(\mathbf{x})\mathbf{k}(\mathbf{x}) \cdot \nabla p(\mathbf{x})) = f \quad \text{in} \quad \Omega, \\ p = 0 \quad \text{on} \quad \partial\Omega, \end{aligned} \tag{4.2.1}$$

The harmonic coordinates G associated with (4.2.1) should satisfy

$$\begin{aligned} \nabla \cdot (\lambda \mathbf{k} \nabla G) = 0 \quad \text{in} \quad \Omega, \\ G(\mathbf{x}) = x \quad \text{on} \quad \partial\Omega. \end{aligned} \tag{4.2.2}$$

By (4.2.2) we mean that G is a n-dimensional vector field $G(x) = (G_1(x) \dots G_n(x))$. It implies that the harmonic coordinates G need to be updated when λ changes, and hence the numerical upscaling methods based on the harmonic coordinates are not attractive in practice. However, the following theorem shows that the compensation phenomena still holds by using the $\mathbf{k}$-harmonic coordinates F, instead of G, when λ is a positive $C^1(\bar{\Omega})$-function. Based on this fact, we develop a multiscale finite element independent of λ to solve (4.2.1) repeatedly.

Theorem 4.2.1. *Let p be the solution of (4.2.1), λ be a positive $C^1(\bar{\Omega})$-function and F be the $\mathbf{k}$-harmonic coordinates defined by (4.1.2). Assume σ is stable and F is an automorphism on Ω, then there exist constants $r > 2$ and $C > 0$ such that $p \circ F^{-1} \in W_0^{2,r}(\Omega)$ and*

$$\|p \circ F^{-1}\|_{W_0^{2,r}(\Omega)} \leq C\|f\|_{L_r(\Omega)}. \tag{4.2.3}$$

The constant r depends on n, Ω, $\lambda_{\max}(\mathbf{k})$, $\lambda_{\min}(\mathbf{k})$, and β_σ. The constant C depends on the constants above and λ, $\|(Trace(\sigma))^{\frac{n}{2p}-1}\|_{L^r(\Omega)}$.

Proof. Since p satisfies the elliptic equation in divergence form (4.2.1), p is in $W_0^{1,r}(\Omega)$ and $\| p \|_{W_0^{1,r}(\Omega)} \leq C\|f\|_{L^r(\Omega)}$ by Meyers Theorem [8]. As λ is a positive $C^1(\bar{\Omega})$-function, we can rewrite (4.2.1) in the following equivalent form

$$\begin{aligned} -\nabla \cdot (\mathbf{k}\nabla p) &= \tilde{f} \quad \text{in} \quad \Omega, \\ p &= 0 \quad \text{on} \quad \partial\Omega, \end{aligned} \tag{4.2.4}$$

where $\hat{f} = \frac{1}{\lambda}(f + \mathbf{k}\nabla\lambda \cdot \nabla p)$. From the assumption that $\mathbf{k}$ and $\nabla\lambda$ are uniformly bounded, we obtain

$$\|\hat{f}\|_{L_r(\Omega)} = \|\frac{1}{\lambda}(f + \mathbf{k}\nabla\lambda \cdot \nabla p)\|_{L_r(\Omega)} \leq C\|f\|_{L_r(\Omega)}. \tag{4.2.5}$$

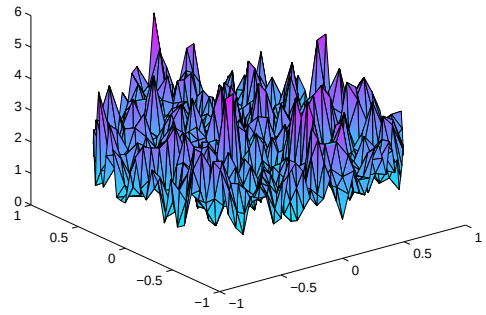
The required result follows by applying Theorem 4.1.1 with (4.2.5). □

Remark 4.2.2. This compensation phenomenon can also be observed numerically. Consider an example: the domain Ω is the unit circle in 2D, $\mathbf{k}$ is a trigonometric multiscale function and λ is $3 + \sin(2\pi x) + \cos(2\pi y)$. In Figure 4.1, we see that both $\nabla(p \circ F^{-1})$ and $\nabla(p \circ G^{-1})$ are smoother than ∇p, and there is not much difference between $\nabla(p \circ F^{-1})$ and $\nabla(p \circ G^{-1})$.

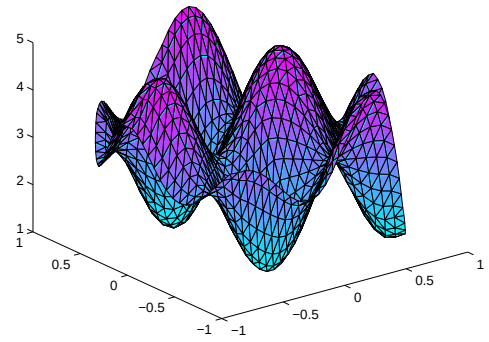
Remark 4.2.3. The smoothness of λ is a necessary condition for Theorem 4.2.1. If λ is not smooth (for example, highly oscillatory), then $p \circ G^{-1}$ may no longer be in $C^{1,\alpha}(\Omega)$ but $p \circ F^{-1}$ is still. See Figure 4.2.

4.2.1 Multiscale Method Using Harmonic Coordinates

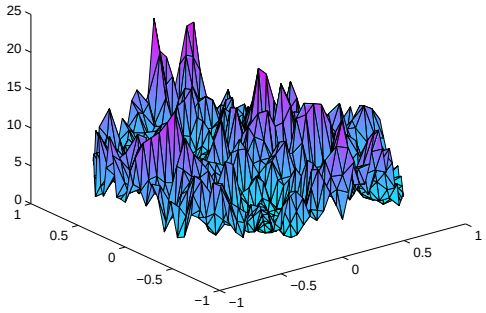
According to Theorem 4.2.1, whatever the choice of h and λ, solutions to (4.2.1) live in the neighborhood of a functional space correlated to F of dimension n. This property can be used in designing numerical methods. In fact, the $\mathbf{k}$-harmonic coordinates can be viewed as single-phase flow solutions with particular boundary conditions. In industry single-phase flow solutions have been widely used in designing multiscale methods for upscaling in two-phase flow problem because of their efficiency and easy implementation. However, there is still a lack of understanding as to why these methods succeed in some cases. In this subsection, we propose a multiscale finite element method using harmonic coordinates (single-phase flow solutions) and provide a rigorous justification of the convergence rate. For simplicity, we describe our method



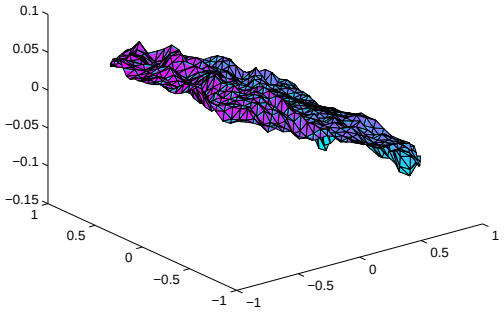
(a) $\mathbf{k}$ is a trigonometric multiscale function.



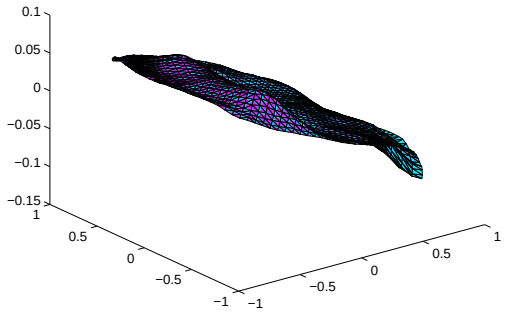
(b) $\lambda = 3 + \sin(2\pi x) + \cos(2\pi y)$



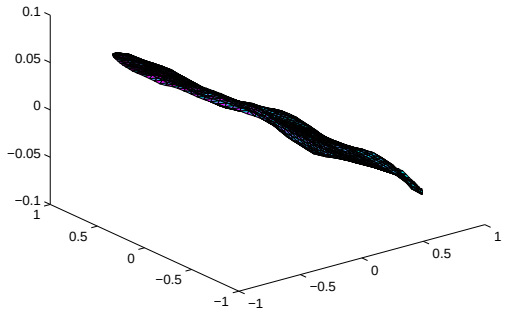
(c) $\lambda\,\mathbf{k}$



(d) The first component of ∇p.

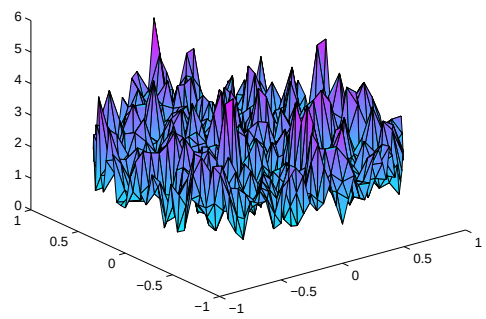


(e) The first component of $\nabla(p \circ F^{-1})$.

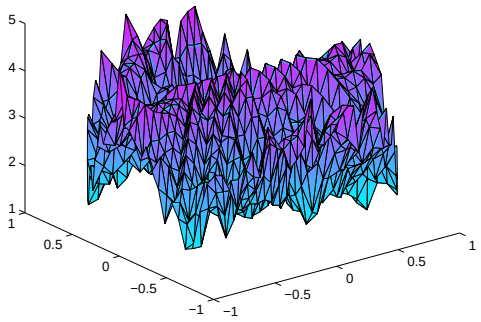


(f) The first component of $\nabla(p \circ G^{-1})$.

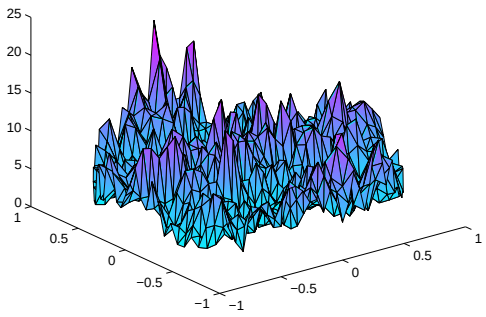
Figure 4.1: Comparison of the gradients of p, $p \circ F^{-1}$ and $p \circ G^{-1}$ when λ is smooth.



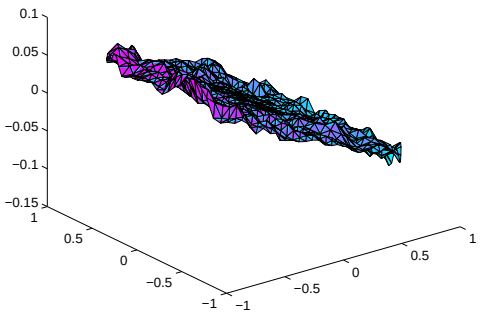
(a) $\mathbf{k}$ is a trigonometric multiscale function.



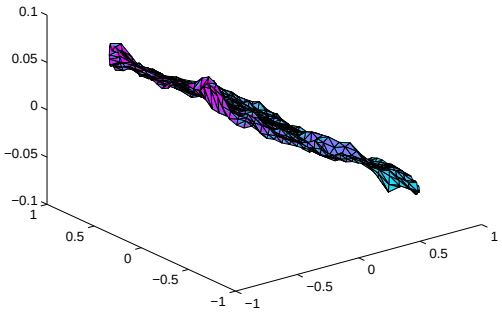
(b) $\lambda = 3 + \sin(2000\pi x) + \cos(2\pi y)$



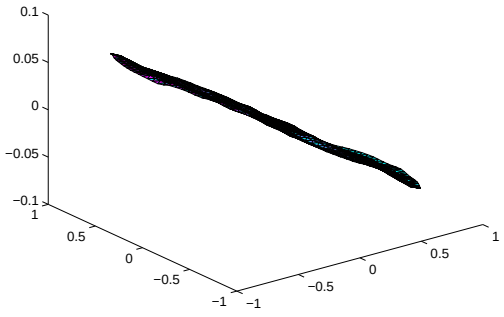
(c) $\lambda\,\mathbf{k}$



(d) The first component of ∇p.



(e) The first component of $\nabla(p \circ F^{-1})$.



(f) The first component of $\nabla(p \circ G^{-1})$.

Figure 4.2: From (e) and (f) we see that the $\mathbf{k}$-harmonic metric fails to adapt the oscillation of $\lambda\mathbf{k}$ when λ is highly oscillatory but $p \circ G^{-1}$ is still smooth.

in 2D only. The method can be easily extended to higher dimensional cases. Our idea is to solve (4.2.1) in the **k**-harmonic coordinates with the standard finite element method.

Let $\mathcal{T}_h$ be a quasi-uniform triangulation on Ω, where h is the maximal length of the mesh. Denote by γ the maximum over the elements τ in $\mathcal{T}_h$ of the ratio between the radius of the smallest ball containing τ and the largest ball inscribed in τ. For a suitable index set $\mathscr{I}_h(\Omega)$, let $\mathcal{N}_h(\Omega)$ denote the nodes of the mesh $\mathcal{T}_h$ which lie in D. Let X_h be the standard finite element space and $\psi_i \in X_h$ be the usual piecewise linear hat functions associated to $x_i \in \mathcal{N}_h(\Omega)$. For each $i \in \mathscr{I}_h(\Omega)$, we construct nodal basis functions Φ_i^{MS} by

$$\Phi_i^{\mathrm{MS}} = \psi_i \circ F. \tag{4.2.6}$$

Define the multiscale finite element space $\mathcal{V}_h^{\mathrm{MS}} = \mathrm{span}\{\Phi_i^{\mathrm{MS}}\}$. Then the multiscale finite element solution $p_h^{\mathrm{MS}} \in \mathcal{V}_h^{\mathrm{MS}}$ satisfies

$$\mathbf{k}(p_h^{\mathrm{MS}}, v_h^{\mathrm{MS}}) \;=\; (f, v_h^{\mathrm{MS}})_{L^2(\Omega)}, \quad \text{for all } v_h^{\mathrm{MS}} \in \mathcal{V}_h^{\mathrm{MS}}, \tag{4.2.7}$$

where $\mathbf{k}(p_h^{\mathrm{MS}}, v_h^{\mathrm{MS}}) = \int_\Omega \mathbf{k}(\mathbf{x}) \nabla p_h^{\mathrm{MS}}(\mathbf{x}) \cdot \nabla v_h^{\mathrm{MS}}(\mathbf{x}) d\mathbf{x}$.

In the standard MsFEM methods, the meshes are determined by users and usually are regular shapes in the computational domain. The basis functions are then computed locally on the meshes. Conversely, the meshes of our method are regular (quasi-uniform triangles) in the harmonic coordinates (physical domain), but the basis function Φ_i^{MS} can be supported on highly distorted and non-local domains since $\mathrm{support}(\Phi_i^{\mathrm{MS}}) = F^{-1}(\mathrm{support}(\psi_i))$. This concept is similar to flow-based coordinate systems [40]. The advantage of using irregular meshes is that the solution p is smoother in the harmonic coordinates. We apply the regularity estimate (4.2.3) to prove the following error estimates.

Theorem 4.2.4. *Assume λ is a positive $C^1(\Omega)$-function and σ satisfies*

$$0 < \lambda_{\min}(\sigma) \leq \lambda_{\max}(\sigma) < \infty.$$

Then there exist constants $r > 2$, $C > 0$ such that

$$|p - p_h^{\mathrm{MS}}|_{H^1(\Omega)} \leq Ch \|f\|_{L_r(\Omega)}. \tag{4.2.8}$$

Furthermore we have the L_2 error estimate

$$\|p - p_h^{\mathrm{MS}}\|_{L_2(\Omega)} \leq Ch^2 \|f\|_{L_r(\Omega)}. \tag{4.2.9}$$

The constant C depends on γ, Ω, $\lambda_{\max}(\mathbf{k})$, $\lambda_{\min}(\mathbf{k})$, μ_σ.

Proof. Since the multiscale finite element space $\mathcal{V}_h^{\mathrm{MS}}$ is contained in $H_0^1(\Omega)$, (4.2.7) gives us the usual optimality estimate

$$|p - p_h^{\mathrm{MS}}|_{H^1(\Omega),\mathbf{k}} \leq \inf_{v_h^{\mathrm{MS}} \in \mathcal{V}_h^{\mathrm{MS}}} |p - v_h^{\mathrm{MS}}|_{H^1(\Omega),\mathbf{k}}, \tag{4.2.10}$$

where $|v|_{H^1(\Omega),\mathbf{k}}$ is the energy norm defined by $\sqrt{\mathbf{k}(v,v)}$. By uniform ellipticity of $\mathbf{k}$, we have $\lambda_{\min}(\mathbf{k}) > 0$ and

$$|p - p_h^{\mathrm{MS}}|^2_{H^1(\Omega)} \leq \frac{1}{\lambda_{\min}(\mathbf{k})} |p - p_h^{\mathrm{MS}}|^2_{H^1(\Omega),\mathbf{k}} \leq \frac{1}{\lambda_{\min}(\mathbf{k})} \inf_{v_h^{\mathrm{MS}} \in \mathcal{V}_h^{\mathrm{MS}}} |p - v_h^{\mathrm{MS}}|_{H^1(\Omega),\mathbf{k}} \tag{4.2.11}$$

Using the change of variable $\mathbf{y} = F(\mathbf{x})$, we obtain

$$\begin{aligned} \inf_{v_h^{\mathrm{MS}} \in \mathcal{V}_h^{\mathrm{MS}}} |p - v_h^{\mathrm{MS}}|^2_{H^1(\Omega),\mathbf{k}} &= \inf_{v_h \in X_h} |p \circ F^{-1} - v_h|^2_{H^1(\Omega),Q} \\ &\leq \lambda_{\max}(Q) \inf_{v_h \in X_h} |p \circ F^{-1} - v_h|^2_{H^1(\Omega)} \\ &\leq C \inf_{v_h \in X_h} |p \circ F^{-1} - v_h|^2_{H^1(\Omega)}, \end{aligned} \tag{4.2.12}$$

where Q is defined by (4.1.5) and can be easily bounded by

$$\lambda_{\max}(Q) \leq (\lambda_{\max}(\mathbf{k})\lambda_{\min}(\mathbf{k}))^{\frac{1}{2}} \mu_\sigma. \tag{4.2.13}$$

Notice that $p \circ F^{-1} \in H^2(\Omega)$ and $\|p \circ F^{-1}\|_{H^2(\Omega)} \leq C\|f\|_{L_r(\Omega)}$ for some $r > 2$ from Theorem 4.2.1. Combining these facts with the standard approximation properties of the hat functions (see for instance [20]) yields

$$\inf_{v_h \in X_h} |p \circ F^{-1} - v_h|_{H^1(\Omega)} \leq Ch \|p \circ F^{-1}\|_{H^2(\Omega)} \leq Ch \|f\|_{L_r(\Omega)}. \tag{4.2.14}$$

The H_1 error estimate (4.2.8) follows straightforwardly from (4.2.11), (4.2.12) and (4.2.14). The L_2 error estimate (4.2.9) can be obtained directly from the Aubin-Nitsche duality argument.[1] □

[1] We refer the reader to Theorem 5.7.6 of [20] for more detail about the duality argument.

4.3 Numerical Experiments

In this section, we perform a numerical experiment to verify the convergence rates established in Theorem 4.2.4. We consider the weak form of the Dirichlet boundary value problem:

$$-\nabla \cdot (\lambda \mathbf{k} \nabla p) = f, \quad \text{in} \quad \Omega, \qquad \text{with} \quad p = 0, \quad \text{on} \quad \partial\Omega\ , \tag{4.3.1}$$

for different λ and $\mathbf{k}$. The computational domain Ω is the unit disk in dimension two and the source term f is the constant function 1. Since an analytical solution is unknown for general λ and $\mathbf{k}$, the equation (4.3.1) is solved on a fine tessellation characterized by 66049 nodes and 131072 triangles as the reference solution to compute the error for solutions on coarser meshes. The $\mathbf{k}$-harmonic coordinates F are solved on the same fine tessellation to construct the basis functions.

Experiment *Trigonometric multiscale*

This example is extracted from [76, 78] as a problem without scale separation:

$$\begin{aligned}\mathbf{k}(x_1, x_2) = \frac{1}{6}\Bigg(&\frac{1.1+\sin(2\pi x_1/\varepsilon_1)}{1.1+\sin(2\pi x_2/\varepsilon_1)} + \frac{1.1+\sin(2\pi x_2/\varepsilon_2)}{1.1+\cos(2\pi x_1/\varepsilon_2)} + \frac{1.1+\cos(2\pi x_1/\varepsilon_3)}{1.1+\sin(2\pi x_2/\varepsilon_3)} \\ &+\frac{1.1+\sin(2\pi x_2/\varepsilon_4)}{1.1+\cos(2\pi x_1/\varepsilon_4)} + \frac{1.1+\cos(2\pi x_1/\varepsilon_5)}{1.1+\sin(2\pi x_2/\varepsilon_5)} + \sin(4x_1^2x_2^2) + 1\Bigg),\end{aligned} \tag{4.3.2}$$

where $\varepsilon_1 = 1/5$, $\varepsilon_2 = 1/13$, $\varepsilon_3 = 1/17$, $\varepsilon_4 = 1/31$, $\varepsilon_5 = 1/65$. We choose two different forms of λ to test our method:

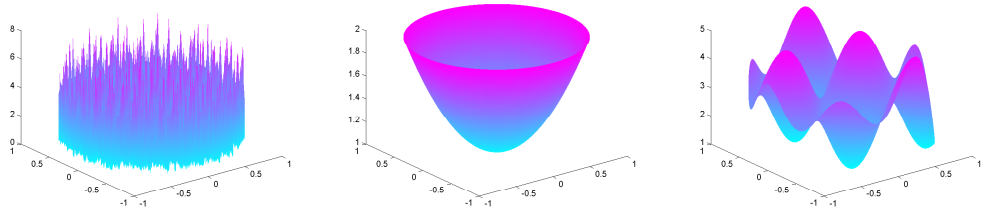
$$\lambda_1(x_1, x_2) \ = \ 1 + x_1^2 + x_2^2$$

and

$$\lambda_2(x_1, x_2, m) \ = \ 3 + \sin(2\pi x_1) + \cos(2\pi x_2).$$

The errors presented in Tables 4.1 and 4.2 show that the method is first order in the H^1 semi-norm and second order in the L_2 norm for both choice of λ, as predicted by the theory.[2]

[2]Throughout, we use least squares fitting to estimate the convergence rates as in Chapter 2.



(a) $\mathbf{k}$ (b) λ_1 (c) λ_2

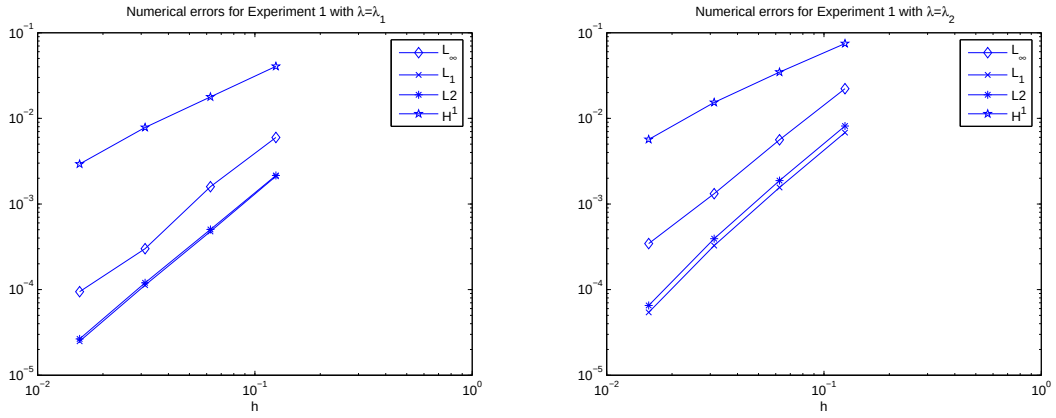
Figure 4.3: Experiment, trigonometric multiscale.

h	L_1	L_2	L_∞	H^1
1/8	2.1094e-3	2.1512e-3	5.9798e-3	4.0551e-2
1/16	4.7713e-4	5.0305e-4	1.5916e-3	1.7797e-2
1/32	1.1284e-4	1.2006e-4	3.0037e-4	7.8346e-3
1/64	2.4952e-5	2.6550e-5	9.4547e-5	2.9255e-3
rate	2.1285	2.1088	2.0354	1.2563

Table 4.1: The L_1, L_2, L_∞ and H^1-norm errors in Experiment with $\lambda = \lambda_1$.

h	L_1	L_2	L_∞	H^1
1/8	6.8181e-3	8.1436e-3	2.2236e-2	7.4690e-2
1/16	1.5581e-3	1.8766e-3	5.6243e-3	3.4630e-2
1/32	3.2637e-4	3.9260e-4	1.3172e-3	1.5312e-2
1/64	5.4230e-5	6.5253e-5	3.4422e-4	5.6696e-3
rate	2.3178	2.3147	2.0134	1.2336

Table 4.2: The L_1, L_2, L_∞ and H^1-norm errors in Experiment with $\lambda = \lambda_2$.



(a) $\lambda = \lambda_1$ (b) $\lambda = \lambda_2$

Figure 4.4: Numerical errors for Experiment with two different λ. The -diamond line is for L_∞ norm error, the -x line is for L_1 norm error, the -asterisk line is for L_2 norm error and the -star is for H^1 norm error.

Chapter 5

Discussions and Future Work

Local boundary condition. In Chapters 2 and 3, we have shown how important the local boundary condition is in designing MsFEM basis functions. Our methods are successful because the local boundary condition adapts the behavior of the exact solutions on the interior boundary. The regularity result of the interface problem serves as a solid foundation for developing methods. There are still many open problems related to constructing optimal local boundary conditions.

Recently Babuška et al. [13] proposed the penetration function to measure the effect of the interior boundary condition. The penetration function is used to assess the accuracy of global-local approaches for recovering local solution from coarse grained solutions. We can apply this technique to examine the optimality of our methods.

Global information. In Chapters 3 and 4, we used global solutions to construct the basis function. More and more research [24, 23, 78, 57, 30] demonstrates the importance of using global information. For example, homogenization without scale separation is very difficult to approach. The metric based up-scaling method [78] provides a solution for this kind of problem by using harmonic solutions. However, it is more expensive to construct the basis functions by using global solutions than traditional MsFEM methods.

One potential way to reduce the cost of computation of global solutions is to design adaptive coarse meshes using the **k**-harmonic transformation, and then construct the MsFEM basis functions locally on these coarse meshes with suitable boundary condition. This enables us to capture the global connectivity of the solution if the heterogeneous media do not have scale separation and the small scale solution has a long range correlation. Since the **k**-harmonic transformation is used for computing the adaptive meshes, we only solve it on a coarse grid. The idea is to combine the

advantages of adaptive mesh techniques and multiscale finite element methods.

5.1 Future Work

Following upon our with multiscale problems, we plan to further develop this research area by studying the following problems.

- **Interface Problems with Many Inclusions and Close Distance**

  In [29], we assume the inclusions are separated from each other with order h distance for analysis. We observe numerically that the solution in Experiment 3 in Section 2.5 becomes more singular as the distance between pairs of inclusions decreases. The exact relation between the regularity of the solution and the distance is unknown. Our method can be applied to such problems, but the hidden constant may blow up as the distance tends to zero. Adjustments are needed to handle the singularity of the solution.

  From the homogenization theory point of view, another challenging problem is to solve a domain containing many inclusions of size ε with distance ρ between neighboring inclusions, where ε and ρ tend to 0 simultaneously. This is an interesting problem especially when the inclusions are not periodic located. Strouboulis et al [86] used the generalized FEM (GFEM) to solve the Laplace equation in a square domain minus 597 voids. This problem can be viewed as the limiting case for many inclusions with conductivities tending to zero.

  We are currently investigating the dependence of the regularity of the solution on ε and ρ with Professors Babych, Kamotski and Smyshlyaev of the University of Bath, UK. Based on this understanding, we can then design suitable boundary condition for basis functions. The numerical solutions should provide a good approximation for the Laplace equation with voids.

- **3-Dimensional Interface Problems**

  The regularity of the solution of the 3D interface problem is similar to that of the solution in the 2D problems (see [53]). However, the geometry in the 3D problem is much more complicated than that of the 2D problems. The intersection of the interface and the boundary of the element is a curve, instead of two points in 2D. To design an appropriate boundary condition for the 3D basis function, we

propose projecting the elliptic equation to the face of the coarse grid elements, which is a 2D plane, and using our 2D method to design the boundary condition.

- **Homogenization of the pressure equation with piecewise smooth mobility**

  In Chapter 4, we proposed a multiscale method to solve the pressure equation when λ is smooth. However, from numerical simulations, we see the saturation S of two-phase flow problems is usually at most piecewise smooth, and hence $\lambda(S)$ is not C^1 on Ω. The next step is to consider the case where λ is a piecewise smooth function, or a piecewise constant function. More precisely, we assume λ has a jump discontinuity along a curve Γ and is C^1 in the sub-domains Ω_1 and Ω_2, where $\Omega = \Omega_1 \bigcup \Omega_2 \bigcup \Gamma$ and $\Gamma = \partial\Omega_1 \bigcap \partial\Omega_2$.

  We propose to use the $\mathbf{k}$-harmonic coordinates F to remove the multiscale effect caused by $\mathbf{k}$ as we have done in Chapter 4. We shall show that $u \circ F^{-1}$ is Hölder continuous on each sub-domain $F(\Omega_1)$ and $F(\Omega_2)$. The transferred equation becomes an interface problem (in harmonic coordinates) which can be solved by the immersed finite element method [68] or the multiscale method we introduced in Chapter 2. Thus we define the multiscale finite element space $\mathcal{V}_h$ as

$$\mathcal{V}_h = \text{span}\{\varphi \circ F : \varphi \in W_h\}, \tag{5.1.1}$$

  where W_h is the immersed finite element space or the multiscale finite element space introduced in Chapter 2.

  The difficulty in applying the proposing method is the smoothness of the interface. To use the IFE or our multiscale method, the interface must be C^1. However, the harmonic coordinates F can be at most C^α and so is $F(\Gamma)$. A new analysis technique must be introduced to overcome this issue.

Appendix A

Proofs of Technical Theorems and Lemmas

A.1 Extension Theorem and Poincaré Inequality on an Annulus

Now we prove some technical lemmas used in Chapter 2.

Lemma A.1.1. Extension theorem on an annulus. *Consider an equilateral triangle denoted τ, with side 1 depicted in Fig A.1, and let τ^δ be the closed annulus of uniform width δ bounded by Γ_1 and Γ_2 as pictured. Let $\Gamma := \Gamma_1 \cup \Gamma_2$ and let $v \in H^{1/2}(\Gamma)$. Then v has an extension $\widetilde{v} \in H^1(\tau^\delta)$ such that*

$$|\widetilde{v}|_{H^1(\tau^\delta)} \leq C\ \delta^{-1} \|v\|_{H^{1/2}(\Gamma)}, \tag{A.1.1}$$

where C is a generic constant independent of δ and v.

Proof. For this proof only, we use the notation $A \lesssim B$ to mean that $A \leq CB$ with C independent of δ and u. We use standard tools for Lipschitz domains (as found for example in [75]). Let $\{W_i\}_{i=1}^N$ be an overlapping open covering of τ^δ and let $\{\phi_i\}_{i=1}^N$ be a corresponding partition of unity with the properties: (i) Each W_i is the intersection of τ^δ with an open ball of diameter $\lesssim \delta$ and either $W_i \cap \Gamma_1 = \emptyset$ or $W_i \cap \Gamma_2 = \emptyset$; (ii) $\|\nabla \phi_i\|_{L^\infty(\tau)} \lesssim \delta^{-1}$; and (iii) each W_i has nonempty intersection with at most m W_j for some number m independent of δ.

Define $\sigma_i = W_i \cap \Gamma$. By property (i), σ_i is either a straight line segment or a corner segment of Γ. Therefore there exists a bijective H^1 map $F_i : \mathbb{R}^2 \to \mathbb{R}^2$ such that $\widetilde{\sigma}_i := F_i(\sigma_i)$ is a subinterval of the x-axis in $\mathbb{R}^2$ and $F_i(W_i)$ is a bounded subset of

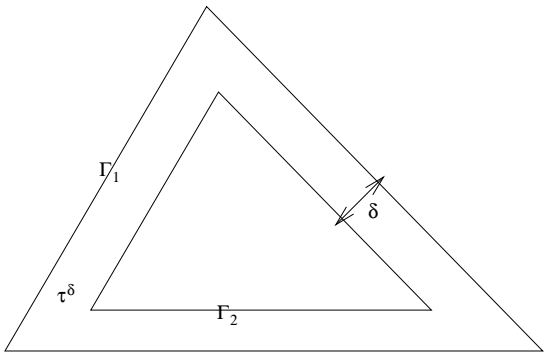


Figure A.1: The graph of the triangular strip.

the upper half plane $\mathbb{R}^2_+$. Since $v \circ F_i^{-1} \in H^{1/2}(\widetilde{\sigma}_i)$ and $\widetilde{\sigma}_i$ is an open subset of $\mathbb{R}$, by the Extension Theorem (e.g. [75, Theorem A.4]), we can find an extension function w_i on $\mathbb{R}$ such that $w_i|_{\widetilde{\sigma}_i} = v \circ F_i^{-1}$ and

$$\|w_i\|_{H^{1/2}(\mathbb{R})} \lesssim \|v \circ F_i^{-1}\|_{H^{1/2}(\widetilde{\sigma}_i)} \lesssim \|v\|_{H^{1/2}(\sigma_i)}.$$

Then, by the (inverse) Trace Theorem on a half-space (e.g. [75, Lemma 3.36]), we can extend w_i to a function $\widetilde{w}_i$ on the upper half plane $\mathbb{R}^2_+$ such that $\widetilde{w}_i \in H^1(\mathbb{R}^2_+)$, $\widetilde{w}_i(x,0) = w_i(x)$, $x \in \mathbb{R}$, and

$$\|\widetilde{w}_i\|_{H^1(\mathbb{R}^2_+)} \lesssim \|w_i\|_{H^{1/2}(\mathbb{R})} \lesssim \|v\|_{H^{1/2}(\sigma_i)}. \tag{A.1.2}$$

Now define $\widetilde{v}_i = \widetilde{w}_i \circ F_i$ and $\widetilde{v} = \sum_{i=1}^N \phi_i \widetilde{v}_i$. Then $\widetilde{v} = v$ on Γ and, using property (ii), (iii) and (A.1.2), we obtain

$$\begin{aligned} |\tilde{v}|^2_{H^1(\tau^\delta)} &\leq \sum_{i=1}^N |\tilde{v}|^2_{H^1(W_i)} \lesssim \sum_{i=1}^N |\phi_i \tilde{v}_i|^2_{H^1(W_i)} \lesssim \delta^{-2} \sum_{i=1}^N \|\tilde{v}_i\|^2_{H^1(W_i)} \\ &\lesssim \delta^{-2} \sum_{i=1}^N \|\tilde{w}_i\|^2_{H^1(\mathbb{R}^2_+)} \lesssim \delta^{-2} \sum_{i=1}^N \|v\|^2_{H^{1/2}(\sigma_i)} \lesssim \delta^{-2} \|v\|^2_{H^{1/2}(\Gamma)}, \end{aligned}$$

as required.

□

Lemma A.1.2. Poincaré inequality on an annulus. *Let τ^δ, δ and $\Gamma = \Gamma_1 \cup \Gamma_2$ be as in Lemma A.1.1. Then for all $u \in H^1(\tau^\delta)$, there exists a constant γ such that*

$$\|u - \gamma\|_{L_2(\tau^\delta)} \lesssim |u|_{H^1(\tau^\delta)} ,$$

where the hidden constant is independent of δ.

Proof. Surprisingly there seems to be no good source for a proof of this result. However a proof may be easily constructed by first writing down the estimate

$$\|u - \gamma\|^2_{L_2(\tau^\delta)} \lesssim \delta \|u - \gamma\|^2_{L^2(\Gamma_2)} + \delta^2 |u|^2_{H^1(\tau^\delta)} .$$

This may be found as equation (3.15) in [88] and can then be combined with the estimate (proven in [79, Lemma 4.3]) that there exists a choice of γ which ensures that

$$\|u - \gamma\|^2_{L^2(\Gamma)} \lesssim \delta^{-1} |u|^2_{H^1(\tau^\delta)} .$$

To be precise, γ may be chosen as the average of u over any of the edges of τ. These two estimates imply the required result. □

A.2 Proof of Theorem 2.2.5

Theorem 2.2.5 *(Babych, Kamotski, Smyshlyaev) Let Ω be either a smooth $\mathcal{C}^\infty$ bounded domain in $\mathbb{R}^2$ or a bounded convex polygon, containing inclusions Ω_i, $i = 1, 2, \ldots, m$, each having a $\mathcal{C}^\infty$ boundary, and define $\Omega_0 = \Omega \setminus \cup_{i=1}^m \overline{\Omega}_i$ as described in Section 2.1. Consider problem* (2.1.2) *and assume that either Case I or Case II (*(2.1.5) *or* (2.1.6)*) holds. Additionally, let $\widetilde{\Gamma}$ denote any closed $\mathcal{C}^\infty$ contour in Ω_0 which encloses all the Ω_i and let $\widetilde{\Omega}_0$ be the domain with boundary $\Gamma \cup \widetilde{\Gamma}$ (see Fig. A.2 for an illustration in the case $m = 1$). Then we have*

$$|u|_{H^{s+2}(\Omega_i)} \lesssim \frac{1}{\alpha_i} \|f\|_{H^s(\Omega)}, \quad \text{for all} \quad s \geq 0 \ , \ i = 1, 2, \ldots, m \ . \tag{A.2.1}$$

Moreover

$$|u|_{H^2(\Omega_0)} \lesssim \frac{1}{\alpha_0} \|f\|_{L_2(\Omega)}, \tag{A.2.2}$$

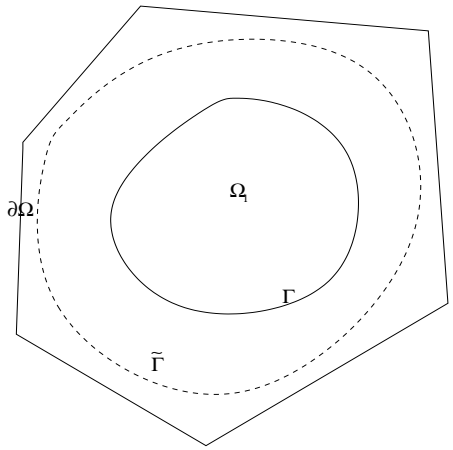


Figure A.2: A domain Ω in the case $m = 1$. The domain Ω_0 is bounded by Γ and $\partial\Omega$ while $\widetilde{\Omega}_0$ is bounded by Γ and $\tilde{\Gamma}$.

and

$$|u|_{H^{2+s}(\widetilde{\Omega}_0)} \lesssim \frac{1}{\alpha_0} \|f\|_{H^s(\Omega)}, \quad \textit{for all} \quad s \geq 0 \ . \tag{A.2.3}$$

The hidden constants depend on the distance of Γ *from* $\partial\Omega$*.*

Proof. In the proof, we consider only the case when Ω is a convex polygon. The case of smooth $\partial\Omega$ is simpler. Also only the case of one inclusion is considered, although the proof for m inclusions is similar. Thus the geometry which we consider is illustrated in Fig. A.2. Also, we shall consider only Case I (2.1.5), i.e.

$$\widehat{\alpha} = \alpha_1 \to \infty \ , \qquad \alpha_0 = 1 \ , \tag{A.2.4}$$

As we shall explain at the end of the proof, Case II is easy once the proof for Case I is clear.

Thus our required result in Case I is, for all $s \geq 0$,

$$|u|_{H^{s+2}(\Omega_i)} \lesssim \frac{1}{\widehat{\alpha}} \|f\|_{H^s(\Omega)}, \quad i = 1, 2 \ldots, m \ , \tag{A.2.5}$$

$$|u|_{H^2(\Omega_0)} \lesssim \|f\|_{L_2(\Omega)}, \qquad \text{and} \quad |u|_{H^{2+s}(\widetilde{\Omega}_0)} \lesssim \|f\|_{H^s(\Omega)} \ . \tag{A.2.6}$$

The result is clear for all $\widehat{\alpha} \in [1, A]$ for some fixed A so we only have to prove (A.2.5), (A.2.6) for $\widehat{\alpha}$ sufficiently large.

Before beginning the proof, we recall two classical regularity results for elliptic

boundary value problems. Let $s \geq 0$ and let $\phi \in H^{s+3/2}(\Gamma)$. Then

$$\left\{ \begin{array}{rcll} \Delta z & = & \omega & \text{on} \quad \Omega_1 \\ z & = & \phi & \text{on} \quad \Gamma \\ \omega & \in & H^s(\Omega_1) & \end{array} \right\} \quad \Longrightarrow \quad \|z\|_{H^{s+2}(\Omega_1)} \lesssim \|\omega\|_{H^s(\Omega_1)} + \|\phi\|_{H^{s+3/2}(\Gamma)} , \tag{A.2.7}$$

and

$$\left\{ \begin{array}{rcll} \Delta z & = & \omega & \text{on} \quad \Omega_0 \\ z & = & \phi & \text{on} \quad \Gamma \\ z & = & 0 & \text{on} \quad \partial\Omega \\ \omega & \in & H^s(\Omega_0) & \end{array} \right\} \quad \Longrightarrow \quad \left\{ \begin{array}{lll} \|z\|_{H^2(\Omega_0)} & \lesssim & \|\omega\|_{L_2(\Omega_0)} + \|\phi\|_{H^{3/2}(\Gamma)} \\ \quad \text{and} & & \\ \|z\|_{H^{s+2}(\widetilde{\Omega}_0)} & \lesssim & \|\omega\|_{H^s(\Omega_0)} + \|\phi\|_{H^{s+3/2}(\Gamma)} \end{array} \right\} . \tag{A.2.8}$$

A suitable reference for (A.2.7) is [69, §2, Theorem 5.4]. For the first inequality on the right-hand side of (A.2.8), we can construct a proof by first extending ϕ to a function $\phi' \in H^2(\Omega_0)$ which vanishes on $\partial\Omega$ and with $\|\phi'\|_{H^2(\Omega_0)} \lesssim \|\phi\|_{H^{3/2}(\Gamma)}$. Then $\Delta(z - \phi') = \omega - \Delta\phi'$ on Ω_0 and $z - \phi'$ vanishes on $\partial\Omega_0$. So using estimates for elliptic equations on domains with convex corners (for example [47, §3.2]) we obtain the estimate. To obtain the second inequality on the right-hand side of (A.2.8), we use interior regularity results (a suitable reference is [87, Theorem 11.1]) and then the trace theorem to obtain

$$\|z\|_{H^{s+3/2}(\tilde{\Gamma})} \lesssim \|\omega\|_{H^s(\Omega_0)} + \|z\|_{H^1(\Omega_0)} \lesssim \|\omega\|_{H^s(\Omega_0)} + \|\phi\|_{H^{1/2}(\Gamma)} .$$

Again applying regularity estimates on the smooth domain $\widetilde{\Omega}_0$ ([69, §2, Thm 5.4]), we obtain the required estimate.

Now the first step in the proof is to introduce a decomposition of the form

$$u = \widehat{u} + \widetilde{u} , \tag{A.2.9}$$

where $\widehat{u}$ solves independent Dirichlet problems with homogeneous boundary data on each Ω_i:

$$-\alpha_i \Delta \widehat{u} = f \quad \text{on} \quad \Omega_i, \quad \text{with} \quad \widehat{u} = 0 \quad \text{on} \quad \partial\Omega_i \quad i = 0, 1 .$$

Then, from (A.2.7) and (A.2.8) and recalling (A.2.4), we obtain, for all $s \geq 0$,

$$\|\widehat{u}\|_{H^{2+s}(\Omega_1)} \lesssim \frac{1}{\widehat{\alpha}}\|f\|_{H^s(\Omega_1)} , \quad \|\widehat{u}\|_{H^2(\Omega_0)} \lesssim \|f\|_{L_2(\Omega_0)} \quad \text{and} \quad \|\widehat{u}\|_{H^{2+s}(\widetilde{\Omega}_0)} \lesssim \|f\|_{H^s(\Omega_0)} . \tag{A.2.10}$$

Thus $\widehat{u}$ satisfies all the estimates (A.2.5), (A.2.6) and the remainder of the proof is concerned with obtaining the same estimates for $\widetilde{u}$. Since $\widetilde{u} = u - \widehat{u}$, it follows that

$$\Delta\widetilde{u} = 0 \quad \text{on} \quad \Omega_i, \quad i = 0, 1 , \quad \text{and} \quad \widetilde{u} = 0 \quad \text{on} \quad \partial\Omega . \tag{A.2.11}$$

Since $\widetilde{u}$ is continuous across Γ, we can define $\widetilde{v} = \widetilde{u}|_\Gamma$.

For any suitably smooth v defined on Ω, we let $\partial v_i/\partial n$ denote the normal derivative of v evaluated on Γ, with value taken from within Ω_i, $i = 0, 1$. (The normal direction is fixed as outward from Ω_1.) Then the usual jump relation for the solution u of the interface problem (2.1.2) reads: $(\partial u_0/\partial n) - \widehat{\alpha}(\partial u_1/\partial n) = 0$, which immediately implies that the function $\widetilde{u}$ satisfies the following equation on Γ:

$$\frac{\partial\widetilde{u}_0}{\partial n} - \widehat{\alpha}\frac{\partial\widetilde{u}_1}{\partial n} \;=\; G \;:=\; \widehat{\alpha}\frac{\partial\widehat{u}_1}{\partial n} - \frac{\partial\widehat{u}_0}{\partial n} . \tag{A.2.12}$$

This may be readily written:

$$(\mathcal{N}_0 - \widehat{\alpha}\mathcal{N}_1)\widetilde{v} \;=\; G , \tag{A.2.13}$$

with $\mathcal{N}_i$ denoting appropriate Dirichlet to Neumann maps on Ω_i (taking $\widetilde{v}$ as Dirichlet data on Γ and using homogeneous Dirichlet data on $\partial\Omega$).

To analyze (A.2.13) as $\widehat{\alpha} \to \infty$, we might consider scaling by $\widehat{\alpha}^{-1}$ to obtain a small perturbation of $\mathcal{N}_1$. However, because $\mathcal{N}_1$ has a non-trivial kernel (namely the constant functions on Γ, henceforth denoted $\langle 1\rangle$), we must study the operator $\mathcal{N}_1$ in the orthogonal complement of this space. Thus we introduce

$$\mathcal{P}v = \frac{1}{|\Gamma|}\int_\Gamma v(s)\; ds ,$$

the orthogonal projection from $L_2(\Gamma)$ onto $\langle 1\rangle$ and $(I - \mathcal{P})$, the orthogonal projection onto $L_2(\Gamma)^\perp := \{v \in L_2(\Gamma) : \mathcal{P}v = 0\}$. Then writing

$$\widetilde{v} \;=\; \mathcal{P}\widetilde{v} \;+\; (I - \mathcal{P})\widetilde{v} =: \widetilde{c} + \widetilde{w} ,$$

equation (A.2.13) may be expressed as the system in $\langle 1\rangle \times L_2(\Gamma)^\perp$:

$$\begin{bmatrix} \mathcal{P}(\mathcal{N}_0 - \widehat{\alpha}\mathcal{N}_1)\mathcal{P} & \mathcal{P}(\mathcal{N}_0 - \widehat{\alpha}\mathcal{N}_1)(I-\mathcal{P}) \\ (I-\mathcal{P})(\mathcal{N}_0 - \widehat{\alpha}\mathcal{N}_1)\mathcal{P} & (I-\mathcal{P})(\mathcal{N}_0 - \widehat{\alpha}\mathcal{N}_1)(I-\mathcal{P}) \end{bmatrix} \begin{bmatrix} \widetilde{c} \\ \widetilde{w} \end{bmatrix} = \begin{bmatrix} \mathcal{P}G \\ (I-\mathcal{P})G \end{bmatrix}. \tag{A.2.14}$$

Moreover, since $\mathcal{P}\mathcal{N}_1 = \mathcal{N}_1\mathcal{P}$ are null operators on $L_2(\Gamma)$, (A.2.14) may be rewritten:

$$(\mathbf{P} - \widehat{\alpha}^{-1}\mathbf{Q}) \begin{bmatrix} \widetilde{c} \\ \widehat{\alpha}\widetilde{w} \end{bmatrix} = \begin{bmatrix} \mathcal{P}G \\ (I-\mathcal{P})G \end{bmatrix}, \tag{A.2.15}$$

where

$$\mathbf{P} = \begin{bmatrix} \mathcal{P}\mathcal{N}_0\mathcal{P} & 0 \\ (I-\mathcal{P})\mathcal{N}_0\mathcal{P} & -\mathcal{N}_1 \end{bmatrix} \quad \text{and} \quad \mathbf{Q} = \begin{bmatrix} 0 & \mathcal{P}\mathcal{N}_0(I-\mathcal{P}) \\ 0 & (I-\mathcal{P})\mathcal{N}_0(I-\mathcal{P}) \end{bmatrix}.$$

We next show that $\mathbf{P}$ is invertible on $\langle 1\rangle \times L_2(\Gamma)^\perp$. Note first that $\mathcal{N}_1$ is invertible on $L_2(\Gamma)^\perp$. To analyse $\mathcal{P}\mathcal{N}_0\mathcal{P}$, consider the boundary value problem:

$$\Delta\eta = 0 \quad \text{in} \quad \Omega_0, \quad \text{with} \quad \eta = 1 \quad \text{on} \quad \Gamma, \quad \text{and} \quad \eta = 0 \quad \text{on} \quad \partial\Omega, \tag{A.2.16}$$

which has a unique solution $\eta \in H^2(\Omega_0)$. The linear operator $\mathcal{P}\mathcal{N}_0\mathcal{P}$ operates on $\langle 1\rangle$ as multiplication by the scalar

$$\gamma := \mathcal{P}\left[\frac{\partial\eta}{\partial n}\right] = \frac{1}{|\Gamma|}\int_\Gamma \frac{\partial\eta}{\partial n}\,ds\,,$$

and this scalar does not vanish, since (by (A.2.16)),

$$\gamma\,|\Gamma| = \int_\Gamma \frac{\partial\eta}{\partial n}\,ds = \int_{\partial\Omega_0} \eta\frac{\partial\eta}{\partial n}\,ds = \int_{\Omega_0} \nabla\cdot(\eta\nabla\eta)\,dx = \int_{\Omega_0} |\nabla\eta|^2\,dx > 0\,.$$

Moreover the linear operator $(I-\mathcal{P})\mathcal{N}_0\mathcal{P}$ operates on $\langle 1\rangle$ as multiplication by the function $\rho := (I-\mathcal{P})(\partial\eta/\partial n) = \partial\eta/\partial n - \gamma \in L_2(\Gamma)^\perp$. Hence

$$\mathbf{P} = \begin{bmatrix} \gamma & 0 \\ \rho & -\mathcal{N}_1 \end{bmatrix} \quad \text{and} \quad \mathbf{P}^{-1} = \begin{bmatrix} \gamma^{-1} & 0 \\ \gamma^{-1}\mathcal{N}_1^{-1}\rho & -\mathcal{N}_1^{-1} \end{bmatrix}.$$

Now combining (A.2.7) and (A.2.8) with the Trace Theorem we obtain that $\mathcal{N}_1 : L_2(\Gamma)^\perp \cap H^{s+3/2}(\Gamma) \to L_2(\Gamma)^\perp \cap H^{s+1/2}(\Gamma)$ is a bounded operator and in fact has a

bounded inverse (see, e.g. [69, §2, Th. 5.4]). Moreover $\mathcal{N}_0 : H^{s+3/2}(\Gamma) \to H^{s+1/2}(\Gamma)$ is also bounded and it is straightforward to show that $\mathbf{P}^{-1}\mathbf{Q}$ is a bounded operator on $\langle 1\rangle \times H^{s+3/2}(\Gamma)$ and that

$$\left\| \mathbf{P}^{-1} \begin{bmatrix} \mathcal{P}G \\ (I-\mathcal{P})G \end{bmatrix} \right\|_{\langle 1\rangle \times H^{s+3/2}(\Gamma)} \lesssim \|G\|_{H^{s+1/2}(\Gamma)} . \tag{A.2.17}$$

Hence, considering (A.2.15) for $\widehat{\alpha}$ sufficiently large, we have the estimate

$$\begin{aligned} \max\{|\widetilde{c}| \ , \ \widehat{\alpha}\|\widetilde{w}\|_{H^{s+3/2}(\Gamma)}\} \quad &\lesssim \quad \|G\|_{H^{s+1/2}(\Gamma)} \ \leq \ \widehat{\alpha} \left\| \frac{\partial \widehat{u}_1}{\partial n} \right\|_{H^{s+1/2}(\Gamma)} + \left\| \frac{\partial \widehat{u}_0}{\partial n} \right\|_{H^{s+1/2}(\Gamma)} \\ &\lesssim \quad \widehat{\alpha}\|\widehat{u}\|_{H^{s+2}(\Omega_1)} \ + \ \|\widehat{u}\|_{H^{s+2}(\widetilde{\Omega}_0)} \qquad (A.2.18) \\ &\lesssim \quad \|f\|_{H^s(\Omega)} \ , \qquad (A.2.19) \end{aligned}$$

where the last three estimates are obtained from employing the definition of G in (A.2.12), then the trace theorem and finally (A.2.10).

Now recall that $\widetilde{u}$ is harmonic on Ω_1 and that $\widetilde{u}|_\Gamma \ =: \ \widetilde{v} \ = \ \widetilde{c} + \widetilde{w}$, where $\widetilde{c} \in \mathbb{R}$. Hence, if we define $\widetilde{u}_1$ on Ω_1 by requiring it to be harmonic and to coincide with $\widetilde{w}$ on Γ, we have, by uniqueness, $\widetilde{u} \ = \ \widetilde{c} \ + \ \widetilde{u}_1$ on Ω_1. Thus by using (A.2.7) and then (A.2.18), we have, for all $s \geq 0$,

$$|\widetilde{u}|_{H^{s+2}(\Omega_1)} \ = \ |\widetilde{u}_1|_{H^{s+2}(\Omega_1)} \ \lesssim \ \|\widetilde{w}\|_{H^{s+3/2}(\Gamma)} \ \lesssim \ \frac{1}{\widehat{\alpha}}\|f\|_{H^s(\Omega)} , \tag{A.2.20}$$

Combining (A.2.20) with the first inequality in (A.2.10) (and recalling (A.2.4)) then yields the first required estimate (A.2.5). To obtain (A.2.6), we note that (A.2.18) implies that $\|\widetilde{v}\|_{H^{s+3/2}(\Gamma)} \ \lesssim \|f\|_{H^s(\Omega)}$ and hence the required estimates follow from (A.2.8).

Finally we remark why the result is easier to prove in Case II, i.e.

$$\widehat{\alpha} = \alpha_0 \to \infty , \qquad \alpha_1 = 1 .$$

In this case the analysis of $\widehat{u}$ is unchanged, but in the analysis of $\widetilde{v}$ we obtain, instead of (A.2.13), the equation

$$(\widehat{\alpha}\mathcal{N}_0 - \mathcal{N}_1)\widetilde{v} \ = \ G := \frac{\partial \widehat{u}_1}{\partial n} - \widehat{\alpha}\frac{\partial \widehat{u}_0}{\partial n} .$$

Since $\mathcal{N}_0$ is invertible the estimate for $\widetilde{v}$ can then be obtained by premultiplying this equation by $\widehat{\alpha}^{-1}\mathcal{N}_0^{-1}$ and letting $\widehat{\alpha}$ get sufficiently large, without having to go through the projection procedure leading to the system (A.2.14). □

Remark A.2.1. Here we briefly discuss the case of inhomogeneous Dirichlet conditions. Consider problem (2.1.2) but replace $u \in H_0^1(\Omega)$ by the requirement that $u \in H^1(\Omega)$ with $u = g$ on $\partial\Omega$. For simplicity assume that g is the restriction to $\partial\Omega$ of a function $g \in H^{s+2}(\Omega_0)$ with $g = 0$ on Γ and with s as large as we wish. Assume also that there is only one inclusion Ω_1. Then an analogous argument to that in Theorem 2.2.5 can be carried out, but with $\widehat{u}$ now required to satisfy $\widehat{u} = g$ on $\partial\Omega$, so that, in Case I, by (A.2.8),

$$\|\widehat{u}\|_{H^{2+s}(\Omega_1)} \lesssim \frac{1}{\widehat{\alpha}}\|f\|_{H^s(\Omega)}\,, \quad \text{and}$$

$$\|\widehat{u}\|_{H^2(\Omega_0)} \lesssim \|f\|_{L_2(\Omega_0)} + \|g\|_{H^2(\Omega_0)} \quad \text{and} \quad \|\widehat{u}\|_{H^{2+s}(\widetilde{\Omega}_0)} \lesssim \|f\|_{H^s(\Omega_0)} + \|g\|_{H^{2+s}(\Omega_0)}.$$

The analysis for $\widetilde{u}$ is as in (A.2.18), (A.2.19) and (A.2.20), leading to the estimates for u:

$$|u|_{H^{2+s}(\Omega_1)} \lesssim \frac{1}{\widehat{\alpha}}\left[\|f\|_{H^s(\Omega)} + \|g\|_{H^{s+2}(\Omega_0)}\right],$$

$$|u|_{H^2(\Omega_0)} \lesssim \|f\|_{L_2(\Omega)} + \|g\|_{H^2(\Omega_0)} \quad \text{and} \quad |u|_{H^{2+s}(\widetilde{\Omega})} \lesssim \|f\|_{H^s(\Omega)} + \|g\|_{H^{2+s}(\Omega_0)}.$$

Thus the higher seminorms of u in the inclusions still decay as $\widehat{\alpha} \to \infty$. However in Case II the seminorms on Ω_0 and $\widetilde{\Omega}_0$ do not necessarily decay and the best estimate for general g is

$$|u|_{H^{2+s}(\Omega_1)} \lesssim \|f\|_{H^s(\Omega)} + \|g\|_{H^{s+2}(\Omega_0)}\,,$$

$$|u|_{H^2(\Omega_0)} \lesssim \frac{1}{\widehat{\alpha}}\left\{\|f\|_{L_2(\Omega)} + \widehat{\alpha}\|g\|_{H^2(\Omega_0)}\right\},$$

and

$$|u|_{H^{2+s}(\widetilde{\Omega}_0)} \lesssim \frac{1}{\widehat{\alpha}}\left\{\|f\|_{H^s(\Omega)} + \widehat{\alpha}\|g\|_{H^{2+s}(\Omega_0)}\right\}.$$

Bibliography

[1] J. Aarnes, On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation, *SIAM J. MMS*, **2** (2004), 421-439.

[2] J. Aarnes, V. Kippe, and K.-A. Lie, Mixed multiscale finite elements and streamline methods for reservoir simulation of large geomodels, *Advances in Water Resources*, **28**(3) (2005), 257-271.

[3] L. Adams and Z. L. Li, The immersed interface/multigird methods for interface problems, *SIAM J. Sci. Comput.* **24** (2002), 463-479.

[4] M. Ainsworth, Robust a posteriori error estimates for nonconforming finite element approximation, *SIAM J. Numer. Anal.*, **42** (2005), 2320 - 2341. G.

[5] G. Alessandrini and V. Nesi, *Univalence of σ-harmonic mappings and applications*, Newton Institute preprint NI00006-SMM, 2000.

[6] G. Alessandrini and V. Nesi, Univalents σ-harmonic mappings, *Arch. Rational Mech. Anal.*, **158** (2001) 155-171.

[7] G. Alessandrini and V. Nesi, Univalent σ-harmonic mappings: connections with quasiconformal mappings, *J. Anal. Math.*, **90** (2003), 197-215.

[8] Allaire, *Shape Optimization by the Homogenization Method*, Springer-Verlag, 2001.

[9] T. Arbogast, Implementation of a locally conservative numerical subgrid upscaling scheme for two-phase Darcy flow, *Comput. Geosci.*, **6** (2002), 453-481.

[10] T. Arbogast, G. Pencheva, M. F. Wheeler, and I. Yotov, A multiscale mortar mixed finite element method, *SIAM J. MMS*, **07** (2007), 319-346.

[11] I. Babuška, The finite element method for elliptic equations with discontinuous coefficients. *Computing*, **5**, (1970), 207-213.

[12] I. Babuška, G. Caloz, and J. E. Osborn, Special finite element methods for a class of second order elliptic problems with rough coefficients, *SIAM J. Numer. Anal.*, **31** (1994), 945-981.

[13] I. Babuška, R. Lipton and M. Stuebner, The penetration function and its application to microscale problems, *BIT Numerical Mathematics*, **48** (2008), 167-197.

[14] I. Babuška and J. E. Osborn. Generalized finite element methods: Their performance and their relation to mixed methods. *SIAM J. Numer. Anal.*, **20** (1983), 510-536.

[15] J. W. Barker and S. Thibeau, A critical review of the use of pseudo-relative permeabilities for upscaling, *SPE Res. Eng.*, **12**, (1997), 138-143.

[16] J.W. Barrett and C.M. Elliott, Fitted and unfitted finite element methods for elliptic-equations with smooth interfaces, *IMA J. Numer. Anal*, **7** (1987), 283 - 300.

[17] J. B. Bell and D. L. Marcus. A second-order projection method for variable-density flows, *J. Comput. Phys.*, **101** (1992), 334-348.

[18] A. Bensoussan, J. L. Lions, and G. Papanicolaou. *Asymptotic Analysis for Periodic Structure.* North Holland, Amsterdam, 1978.

[19] C. Bernardi and R. Verfürth, Adaptive finite element methods for elliptic equations with non-smooth coefficients, *Numer. Math.*, **85** (2000), 579-608.

[20] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods.* Springer-Verlag, 2002.

[21] I-L. Chern and Y.-C. Shu, A coupling interface method for elliptic interface problems, *J. Comp. Phys.*, **225**, (2007), 2138-2174.

[22] Z. Chen and J. Zou, Finite element methods and their convergence for elliptic and parabolic interface problems, *Numer. Math.*, **79** (1998), 175-202.

[23] Y. Chen and L. J. Durlofsky, Adaptive coupled local-global upscaling for general flow scenarios in heterogeneous formations, *Transport in Porous Media*, **62** (2006), 157-185.

[24] Y. Chen, L. J. Durlofsky, M. Gerritsen, and X. H. Wen, A coupled local-global upscaling approach for simulating flow in highly heterogeneous formations, *Advances in Water Resources*, **26** (2003), 1041-1060.

[25] Z. Chen and T. Y. Hou, A mixed multiscale finite element method for elliptic problems with oscillating coefficients, *Math. Comp.*, **72** (2002), 541-576 (electronic).

[26] A. J. Chorin, Numerical solution of the Navier-Stokes equations. *Math. Comp.*, **22** (1968), 745-762.

[27] M. Christie, Upscaling for reservoir simulation, *J. Pet. Tech.*, (1996), 1004-1010.

[28] M. Christie and M. Blunt, Tenth spe comparative solution project: A comparison of upscaling techniques, *SPE Reser. Eval. Eng.*, **4** (2001), 308-317.

[29] C.-C. Chu, I.G. Graham, T. Y. Hou, A new multiscale finite element method for high-contrast elliptic interface problems. Submitted.

[30] J. Chu, Y. Efendiev, V. Ginting, and T. Y. Hou, Flow based oversampling technique for multiscale finite element methods , *Advances in Water Resources*, **31** (2008) 599-608.

[31] M. Dryja, On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients, *Computational Methods in Applied Mathematics*, **3** (2003), 76-85.

[32] M.A. Dumett, J.P. Keener, An immersed interface method for solving anisotropic elliptic boundary value problems in three dimensions, *SIAM J. Sci. Comput.*, **25** (2003), 348-367

[33] L. J. Durlofsky, Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media, *Water Resour. Res.*, **27** (1991), 699-708.

[34] L. J. Durlofsky, Coarse scale models of two phase flow in heterogeneous reservoirs: Volume averaged equations and their relationship to the existing upscaling techniques, *Comput. Geosci.*, **2** (1998),‘73-92.

[35] L. J. Durlofsky, Y. Efendiev, and V. Ginting, An adaptive local-global multiscale finite volume element method for two-phase flow simulations, *Advances in Water Resources*, **30** (2007), 576-588

[36] W. E and B. Engquist, The heterogeneous multiscale methods, *Comm. Math. Sci.*, **1**(1) (2003), 87-133.

[37] Y. Efendiev, V. Ginting, T. Hou, and R. Ewing, Accurate multiscale finite element methods for two-phase flow simulations, *Journal of Computational Physics*, **220** (2006), 155-174.

[38] Multiscale finite element methods for nonlinear problems and their applications, *Comm. Math. Sci.*, **2** (2004), 553-589.

[39] Y. Efendiev and T. Y. Hou, *Multiscale Finite Element Methods: Theory and Applications*, Springer-Verlag, 2009.

[40] Y. Efendiev, T. Y. Hou, and T. Strinopoulos, Multiscale Simulations of Porous Media Flows in Flow-Based Coordinate System, *Comput. Geoscience*, **12**, (2008), 257-272.

[41] Y. Efendiev, T. Hou, and X. H. Wu, Convergence of A Nonconforming Multiscale Finite Element Method, *SIAM J. Numer. Anal.*, **37** (2000), 888-910.

[42] R.P. Fedkiw, T. Aslam, B, Merriman, S. Osher, A non-oscillatory Eulerian approach to interfaces in multimaterial flows (the ghost fluid method), *J. Comput. Phys.* **161** (2000), 30-60.

[43] M. Gerritsen and L. Durlofsky, Modeling of fluid flow in oil reservoirs, *Annual Reviews in Fluid Mechanics*, **37** (2005), 211-238.

[44] E. De Giorgi, G-operators and γ- convergence, *Proc. Int. Congr. Math.*, 1175-1191. PWN Polish Scientific Publishers and North-Holland, 1984.

[45] I.G. Graham, P. Lechner and R. Scheichl, Domain Decomposition for Multiscale PDEs, *Numer. Math.*, **106** (2007), 471-510.

[46] I.G. Graham and R. Scheichl, Robust Domain Decomposition Algorithms for Multiscale PDEs, *Numerical Methods for Partial Differential Equations*, **23** (2007), 859-878.

[47] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.

[48] T. Y. Hou, Z.L. Li, S. Osher, H. Zhao, A hybrid method for moving interface problems with application to the Hele-Shaw flow, *J. Comp. Phys.* **134** (1997), 236-252.

[49] S. Hou and X-D. Liu, A numerical method for solving variable coefficient elliptic equation with interfaces, *J. Comp. Phys.* **202** (2005), 411-445.

[50] T. Y. Hou, J. S. Lowengrub, and M. J. Shelley, Removing the stiffness from interfacial flow with surface tension, *J. Comput. Phys.*, **114** (1994), 312-338.

[51] T. Y. Hou and X.-H. Wu, A multiscale finite element method for elliptic problems in composite materials and porous media, *J. Comput. Phys.* **134** (1997), 169-189.

[52] T. Y. Hou, X.-H. Wu and Z. Cai, Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients, *Math. Comp.* **68** (1999), 913-943.

[53] J. Huang and J. Zou, Some new a priori estimates for second-order elliptic and parabolic interface problems, *Journal of Differential Equations* **184** (2002), 570-586.

[54] T. Hughes, G. Feijoo, L. Mazzei, and J. Quincy, The variational multiscale method-a paradigm for computational mechanics, *Comput. Methods Appl. Mech. Engrg*, **166** (1998), 3-24.

[55] P. Jenny, S. H. Lee, and H. Tchelepi, Multi-scale finite volume method for elliptic problems in subsurface flow simulation, *J. Comput. Phys.*, **187** (2003), 47-67.

[56] P. Jenny, S. H. Lee, and H. Tchelepi, Adaptive multi-scale finite volume method for multi-phase flow and transport in porous media, *Multiscale Modeling and Simulation*, **3** (2005), 30-64.

[57] L. Jiang, Y. Efendiev, and V. Ginting, Multiscale methods for parabolic equations with continuum spatial scales, *Discrete Contin. Dyn. Syst. Ser.* **8** (2007), 833-859.

[58] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik. *Homogenization of Differential Operators and Integral Functionals.* Springer-Verlag, 1991.

[59] R. B. Kellogg, Higher order singularities for interface problems. In B. Hubbard, editor, *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, 589-602. Academic Press, New York, 1972.

[60] R. B. Kellogg, Singularities in interface problems. In A.K. Aziz, editor,*Numerical Solution of Partial Differential Equations*, 351-400, Academic Press, New York, 1971.

[61] B.C. Khoo, T.G. Liu and C.W. Wang, The ghost fluid method for compressible gas-water simulation, *J. Comp. Phys.* **204** (2005), 193-221.

[62] B.C. Khoo, T.G. Liu and K.S. Yeo, Ghost fluid method for strong shock impacting on meterial interface, *J. Comp. Phys.* **190** (2003), 651-681.

[63] R. J. LeVeque. Clawpack and Amrclaw – Software for high-resolution Godunov methods. 4-th Intl. Conf. on Wave Propagation, Golden, Colorado, 1998.

[64] R. J. LeVeque and Z.L. Li, The immersed interface method for elliptic equations with discontinuous coefficients and singular sources, *SIAM J. Numer. Anal.*, **31** (1994), 1019–1044.

[65] J. Li, J.M. Melenk, B. Wohlmuth and J. Zou, Optimal convergence of higher order finite element methods for elliptic interface problems ASC Report 13/2008, Institute for Analysis and Scientific Computing, T.U. Vienna (2008).

[66] Z.L. Li, A fast iterative algorithm for elliptic interface problems, *SIAM J. Numer. Anal.* **35** (1998), 230-254

[67] Z.L. Li, K. Ito, Maximum priciple preserving schemes for interface problems with discontinuous coefficients, *SIAM J. Sci. Comput.* **23** (2001), 339-361

[68] Z.L. Li, T. Lin and X. Wu, New Cartesian grid methods for interface problems using the finite element formulation, *Numerische Mathematik* **96** (2003), 61-98.

[69] J.L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*, Springer-Verlag, Berlin, 1972.

[70] W. Littman, G. Stampacchia and H. F. Weinberger, Regular points for elliptic equations with discontinuous coefficients, *Ann. Scuola Norm. Sup. Pisa (3)* **17**, 43-77.

[71] X.D. Liu, R.P. Fedkiw, M.J. Kang, A boundary condition capturing method for Poisson's equation on irregular domains, *J. Comp. Phys.* **160** (2000), 151-178.

[72] X.D. Liu, T. C. Sideris, Convergence of the ghost fluid method for elliptic equations with interfaces, *Math. Comput.* **72** (2003), 1731-1746

[73] A.-M. Matache and C. Schwab, Homogenization via p-FEM for problems with microstructure, *Proceedings of the Fourth International Conference on Spectral and High Order Methods* (ICOSAHOM 1998) (Herzliya), **33**, 2000, 43-59.

[74] A. Maugeri, D. K. Palagachev, and L. G. Softova. *Elliptic and Parabolic Equations with Discontinuous Coefficients*, volume 109 of Mathematical Researchm, Wiley- VCH, 2000.

[75] W. Mclean, Strongly elliptic systems and boundary integral equations, *Cambridge University Press*, Cambridge, UK (1998).

[76] P. Ming and X. Yue, Numerical methods for multiscale elliptic problems, *J. of Comput. Phys.*, **214** (2006), 421-445.

[77] F. Murat and L. Tartar, H-convergence. In R . V. Kohn, editor, *Séminaire d'Analyse Fonctionnelle et Numérique de l'Université d'Alger(1977)*, 1977.

[78] H. Owhadi and L. Zhang, Metric based up-scaling, *Comm. Pure Appl. Math.*, **60** (2006), 675-723.

[79] C. Pechstein and R. Scheichl, Analysis of FETI methods for multiscale PDEs, *Numer. Math.* **111** (2008), 293-333.

[80] C. S. Peskin, Numerical study of blood flow in the heart, *J. Comput. Phys.*, **25** (1977), 220-252.

[81] C. S. Peskin, The immersed boundary method, *Acta Numerica*, **11** (2002), 479-517.

[82] M. Plum and C. Wieners, Optimal a priori estimates for interface problems, *Numer. Math.*, **95** (2003), 735–759.

[83] P. Renard and G. deMarsily, Calculating effective permeability: a review, *Advances in Water Resources*, **20** (1997), 253-278.

[84] G. Sangalli, Capturing small scales in elliptic problems using a residual-free bubbles finite element method, *Multiscale Model. Simul.*, **1** (2003), 485-503 (electronic).

[85] S. Spagnolo, Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche. *Ann. Scuola Norm. Sup. Pisa (3)*, **22** (1968), 571-597, 1968.

[86] T. Strouboulis, L. Zhang and I. Babuška, Assessment of the cost and accuracy of the generalized FEM, *Int. J. Numer. Meth. Engrg.*, **69** (2007), 250-283.

[87] M.E. Taylor, *Partial Differential Equations*, Volume I, Springer-Verlag, Berlin, 1996.

[88] A. Toselli and O. Widlund, *Domain Decomposition Methods Algorithms and Theory*, Springer-Verlag, Berlin-Heidelberg-New York, 2005.

[89] S.O. Unverdi and G. Tryggvason, A front-tracking method for viscous, incompressible, multi-fluid flows, *J. Comput. Phys.*, **100** (1992), 25-37.

[90] M. Vohralík, Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients, preprint, University of Paris VI, 2008.

[91] X. H. Wen and J. J. Gomez-Hernandez, Upscaling hydraulic conductivities in heterogeneous media: An overview, *Journal of Hydrology*, **183** (1996), ix–xxxii.

[92] Y.C. Zhou and G.W. Wei, On the fictitious-domain and interpolation formulations of the matched and boundary (MIB) method, *J. Comp. Phys.*, **219** (2006) 228-246.

[93] Y.C. Zhou, S. Zhao, M. Feig and G.W. Wei, High order matched interface and boundary method for elliptic equations with discontinuous coefficients and singular sources, *J. Comp. Phys.*, **213** (2006), 1-30.